

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 17:14:02 ; Search time 1718 Seconds  
(without alignments)  
564.089 Million cell updates/sec

Title: US-10-672-399-10

Perfect score: 20

Sequence: 1 agggcgctctctgagtagcag 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_to.\*

11: gb\_sts.\*

12: gb\_by.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	1581	6	CQ731079 Sequence
C 2	20	100.0	2088	9	U59269 Human hyalu
C 3	20	100.0	2108	9	D84424 Homo sapien
C 4	20	100.0	2116	6	AR220003 Sequence
C 5	20	100.0	2117	6	AR137038 Sequence
C 6	20	100.0	2119	9	BC035837 Homo sapi
C 7	20	100.0	193986	2	AC130783 Pan trogl
C 8	20	100.0	229155	9	AC018755 Homo sapi
C 9	18.4	92.0	517	4	AB017803 Bos tauri
C 10	17.4	87.0	600	11	BV099316 RPAMSEQO
C 11	17.4	87.0	600	11	BV160504 RPAMSEQO
C 12	17.4	87.0	2095	10	AB097568 Rattus no
C 13	17.4	87.0	2102	6	E13681 DNA encodin
C 14	17.4	87.0	2102	6	E28454 Hyaluronate
C 15	17.4	87.0	2102	6	E30971 Hyaluronic
C 16	17.4	87.0	2102	6	E34326 DNA for gen
C 17	17.4	87.0	2102	10	D82964 Mus musculu
C 18	17.4	87.0	2109	9	AY463695 Papio anu
C 19	17.4	87.0	72955	2	AC108651_3 Continuation (4 of

20	17.4	87.0	185623	2	AC137058	AC137058 Papio anu
21	17.4	87.0	189862	2	AC108317	AC108317 Rattus no
22	17.4	87.0	213263	10	AL593846	AL593846 Mouse DNA
23	17.4	87.0	257325	2	AC116203	AC116203 Rattus no
24	17.4	87.0	293184	2	AC079487	AC079487 Mus muscu
25	16.8	84.0	278	6	AX898222	AX898222 Sequence
26	16.8	84.0	278	6	BD033755	BD033755 Sequence
27	16.8	84.0	313	6	AX886427	AX886427 Sequence
28	16.8	84.0	313	6	BD026037	BD026037 Sequence
29	16.8	84.0	465	9	AF367431	AF367431 Homo sapi
30	16.8	84.0	529	9	AF367432	AF367432 Homo sapi
31	16.8	84.0	536	9	AF367430	AF367430 Homo sapi
32	16.8	84.0	555	11	BV037702	BV037702 S212P6015
33	16.8	84.0	774	9	AF267852	AF267852 Homo sapi
34	16.8	84.0	795	11	BV063966	BV063966 S212P6739
35	16.8	84.0	3111	9	AX056640	AX056640 Homo sapi
36	16.8	84.0	43859	9	AC006133	AC006133 Homo sapi
37	16.8	84.0	45136	2	AC100349	AC100349 Mus muscu
38	16.8	84.0	80065	2	AC016271	AC016271 Homo sapi
39	16.8	84.0	110000	8	CR382131_11	Continuation (12 o
40	16.8	84.0	127109	9	AC106708	AC106708 Homo sapi
41	16.8	84.0	128198	10	AL731701	AL731701 Mouse DNA
42	16.8	84.0	141674	9	AC026671	AC026671 Homo sapi
43	16.8	84.0	142447	10	AL627204	AL627204 Mouse DNA
44	16.8	84.0	144791	2	AC142379	AC142379 Rattus no
45	16.8	84.0	149727	2	AC073808	AC073808 Mus muscu
46	16.8	84.0	150010	9	AC104782	AC104782 Homo sapi
47	16.8	84.0	150354	2	AC133622	AC133622 Rattus no
48	16.8	84.0	155211	2	AC019258	AC019258 Homo sapi
49	16.8	84.0	159431	10	AC128702	AC128702 Mus muscu
50	16.8	84.0	160872	2	AC053534	AC053534 Homo sapi
51	16.8	84.0	172790	2	AC024165	AC024165 Homo sapi
52	16.8	84.0	173643	2	AC119743	AC119743 Homo sapi
53	16.8	84.0	174217	2	AC021408	AC021408 Homo sapi
54	16.8	84.0	180816	2	AC110395	AC110395 Rattus no
55	16.8	84.0	182700	2	AC145747	AC145747 Mus muscu
56	16.8	84.0	183759	2	AC010935	AC010935 Homo sapi
57	16.8	84.0	183865	2	AC130434	AC130434 Homo sapi
58	16.8	84.0	191093	10	AL671765	AL671765 Mouse DNA
59	16.8	84.0	192663	10	AC134406	AC134406 Mus muscu
60	16.8	84.0	194366	10	AC109610	AC109610 Mus muscu
61	16.8	84.0	194454	2	AC012357	AC012357 Homo sapi
62	16.8	84.0	206321	2	AC124131	AC124131 Mus muscu
63	16.8	84.0	206677	10	AC129085	AC129085 Mus muscu
64	16.8	84.0	210172	2	AC121646	AC121646 Rattus no
65	16.8	84.0	210928	2	AC118097	AC118097 Rattus no
66	16.8	84.0	214436	2	AC119510	AC119510 Rattus no
67	16.8	84.0	216139	2	AC132676	AC132676 Rattus no
68	16.8	84.0	217943	2	AC127046	AC127046 Rattus no
69	16.8	84.0	218698	9	AC007126	AC007126 Homo sapi
70	16.8	84.0	220766	2	AC103203	AC103203 Rattus no
71	16.8	84.0	225007	2	AC110885	AC110885 Mus muscu
72	16.8	84.0	225698	10	AC100212	AC100212 Mus muscu
73	16.8	84.0	229118	2	AC114408	AC114408 Mus muscu
74	16.8	84.0	230056	2	AC134124	AC134124 Rattus no
75	16.8	84.0	230569	2	AC108250	AC108250 Rattus no
76	16.8	84.0	236164	10	AC101205	AC101205 Mus muscu
77	16.8	84.0	236344	10	AL583889	AL583889 Mouse DNA
78	16.8	84.0	238608	2	AC095637	AC095637 Rattus no
79	16.8	84.0	238999	2	AC128410	AC128410 Rattus no
80	16.8	84.0	243704	2	AC099277	AC099277 Rattus no
81	16.8	84.0	243981	2	AC111281	AC111281 Rattus no
82	16.8	84.0	252612	2	AC111263	AC111263 Rattus no
83	16.8	84.0	253057	2	AC115253	AC115253 Rattus no
84	16.8	84.0	254197	2	AF389853	AF389853 Mus muscu
85	16.8	84.0	258293	2	AC114379	AC114379 Rattus no
86	16.8	84.0	259757	2	AC130146	AC130146 Rattus no
87	16.8	84.0	268756	2	AC117345	AC117345 Rattus no
88	16.8	84.0	276958	2	AC110636	AC110636 Rattus no
89	16.4	82.0	300	6	BD128919	BD128919 Human gen
90	16.4	82.0	452	6	BD129543	BD129543 Human gen
91	16.4	82.0	1191	6	CQ724801	CQ724801 Sequence
92	16.4	82.0	1611	6	AX747569	AX747569 Sequence

93	16.4	82.0	1611	9	AK092491	AK092491 Homo sapi	c 166	16.4	82.0	241696	2	AC130558	AC130558 Rattus no
94	16.4	82.0	2057	6	BD159624	BD159624 Primer fo	c 167	16.4	82.0	244577	2	AC129706	AC129706 Rattus no
95	16.4	82.0	2057	6	AX862261	AX862261 Sequence	c 168	16.4	82.0	246469	2	AC131647	AC131647 Rattus no
96	16.4	82.0	2057	6	AX822756	AX822756 Homo sapi	c 169	16.4	82.0	249184	2	AC127181	AC127181 Rattus no
97	16.4	82.0	2250	6	AX840058	AX840058 Sequence	c 170	16.4	82.0	253188	2	AC111666	AC111666 Rattus no
98	16.4	82.0	2753	6	BD160551	BD160551 Primer fo	c 171	16.4	82.0	253188	2	AC127877	AC127877 Rattus no
99	16.4	82.0	2753	6	AX883823	AX883823 Sequence	c 172	16.4	82.0	254698	2	AC094849	AC094849 Rattus no
100	16.4	82.0	2753	9	AK024284	AK024284 Homo sapi	c 173	16.4	82.0	255541	2	AC097730	AC097730 Rattus no
101	16.4	82.0	2801	9	BC003555	BC003555 Homo sapi	c 174	16.4	82.0	256899	2	AC097413	AC097413 Rattus no
102	16.4	82.0	15105	6	AX840062	AX840062 Sequence	c 175	16.4	82.0	257097	2	AC096199	AC096199 Rattus no
103	16.4	82.0	70720	2	AC027823	AC027823 Homo sapi	c 176	16.4	82.0	260873	2	AC121675	AC121675 Rattus no
104	16.4	82.0	83078	2	AC079509	AC079509 Mus muscu	c 177	16.4	82.0	263052	2	AC099360	AC099360 Rattus no
105	16.4	82.0	88087	2	AC110708	AC110708 Rattus no	c 178	16.4	82.0	263349	2	AC134358	AC134358 Rattus no
c 106	16.4	82.0	89729	5	EX649321	EX649321 Zebrafish	c 179	16.4	82.0	275589	2	AC105503	AC105503 Rattus no
107	16.4	82.0	107455	9	AL512308	AL512308 Human DNA	c 180	16.4	82.0	280575	2	AC117364	AC117364 Rattus no
108	16.4	82.0	144432	2	AF276983	AF276983 Homo sapi	c 181	16.4	82.0	293757	10	AC095491	AC095491 Rattus no
109	16.4	82.0	154542	2	AC079040	AC079040 Mus muscu	c 182	16.4	82.0	294494	2	AC098385	AC098385 Rattus no
110	16.4	82.0	163399	2	AC079860	AC079860 Homo sapi	c 183	16.4	82.0	297293	2	AC125863	AC125863 Rattus no
111	16.4	82.0	165453	2	AC121860	AC121860 Mus muscu	c 184	16.4	82.0	301888	2	AC094881	AC094881 Rattus no
112	16.4	82.0	168544	2	AC131134	AC131134 Rattus no	c 185	16.4	82.0	304517	2	AC096282	AC096282 Rattus no
113	16.4	82.0	171978	10	AC125138	AC125138 Mus muscu	c 186	16.4	82.0	327883	2	EX571684	EX571684 Homo sapi
114	16.4	82.0	173251	2	AC139823	AC139823 Homo sapi	c 187	16.4	82.0	336984	2	AC113635	AC113635 Rattus no
115	16.4	82.0	175883	9	AP000808	AP000808 Homo sapi	c 188	16.4	80.0	384	9	HS239330	HS239330 Homo sapi
c 116	16.4	82.0	176757	4	AC095025	AC095025 Sus scrofa	c 189	16.4	80.0	63451	9	AL662886	AL662886 Human DNA
c 117	16.4	82.0	177689	10	AC116500	AC116500 Mus muscu	c 190	16.4	80.0	68997	9	AC092861	AC092861 Homo sapi
c 118	16.4	82.0	178382	2	AC139481	AC139481 Homo sapi	c 191	16.4	80.0	84797	8	AC011661	AC011661 genomic s
c 119	16.4	82.0	180005	2	AC136070	AC136070 Rattus no	c 192	16.4	80.0	147009	2	AC084179	AC084179 Homo sapi
120	16.4	82.0	182072	2	AC133183	AC133183 Mus muscu	c 193	16.4	80.0	154467	2	AC022566	AC022566 Homo sapi
121	16.4	82.0	182715	2	AC130031	AC130031 Rattus no	c 194	16.4	80.0	243812	2	AC118510	AC118510 Rattus no
c 122	16.4	82.0	183005	10	AC132570	AC132570 Mus muscu	c 195	15.8	79.0	525	6	AR309640	AR309640 Sequence
c 123	16.4	82.0	184891	2	AC108326	AC108326 Rattus no	c 196	15.8	79.0	525	6	AR430174	AR430174 Sequence
c 124	16.4	82.0	186759	9	AL645608	AL645608 Human DNA	c 197	15.8	79.0	525	6	AX195141	AX195141 Sequence
125	16.4	82.0	186836	9	AC079340	AC079340 Homo sapi	c 198	15.8	79.0	525	6	AX683840	AX683840 Sequence
c 126	16.4	82.0	186848	2	AC145556	AC145556 Mus muscu	c 199	15.8	79.0	737	8	AK064693	AK064693 Oryza sat
c 127	16.4	82.0	190078	2	AC139289	AC139289 Homo sapi	c 200	15.8	79.0	779	6	BD220581	BD220581 Human gen
128	16.4	82.0	190752	2	AC080159	AC080159 Mus muscu	c 201	15.8	79.0	841	10	AF183962	AF183962 Mus muscu
c 129	16.4	82.0	194199	2	AC135287	AC135287 Rattus no	c 202	15.8	79.0	946	10	AF463665	AF463665 Mus muscu
c 130	16.4	82.0	194931	10	AC125063	AC125063 Mus muscu	c 203	15.8	79.0	1094	8	AK111978	AK111978 Oryza sat
c 131	16.4	82.0	196127	2	AC132788	AC132788 Rattus no	c 204	15.8	79.0	1177	8	AF500581	AF500581 Apicarpa
c 132	16.4	82.0	196715	2	AC141600	AC141600 Homo sapi	c 205	15.8	79.0	1505	8	AK059039	AK059039 Oryza sat
c 133	16.4	82.0	198026	2	AC124730	AC124730 Mus muscu	c 206	15.8	79.0	1598	10	BC062643	BC062643 Mus muscu
c 134	16.4	82.0	200491	2	AB107102	AB107102 Homo sapi	c 207	15.8	79.0	1864	5	AY232496	AY232496 Synanceia
c 135	16.4	82.0	200702	2	AC130278	AC130278 Mus muscu	c 208	15.8	79.0	1991	9	BC007698	BC007698 Homo sapi
c 136	16.4	82.0	203291	2	AC124762	AC124762 Mus muscu	c 209	15.8	79.0	2069	6	CQ850809	CQ850809 Sequence
c 137	16.4	82.0	203788	2	AC105965	AC105965 Mus muscu	c 210	15.8	79.0	2069	9	AK127991	AK127991 Homo sapi
c 138	16.4	82.0	205893	2	AC115530	AC115530 Rattus no	c 211	15.8	79.0	2220	14	HSV52K	M21943 Hepesvirus
c 139	16.4	82.0	208868	2	AC115946	AC115946 Mus muscu	c 212	15.8	79.0	2373	4	AF184968	AF184968 Oryctolag
c 140	16.4	82.0	211552	2	AC141271	AC141271 Homo sapi	c 213	15.8	79.0	2561	10	AF199491	AF199491 Mus muscu
c 141	16.4	82.0	212610	2	AC143325	AC143325 Homo sapi	c 214	15.8	79.0	2614	6	AR064012	AR064012 Sequence
c 142	16.4	82.0	214441	5	EX088571	EX088571 Zebrafish	c 215	15.8	79.0	2614	6	BD235630	BD235630 Single ge
c 143	16.4	82.0	214809	2	AC132744	AC132744 Rattus no	c 216	15.8	79.0	2614	6	AR195063	AR195063 Sequence
c 144	16.4	82.0	214868	2	AC139122	AC139122 Rattus no	c 217	15.8	79.0	2614	6	AR212272	AR212272 Sequence
c 145	16.4	82.0	216381	10	AC120138	AC120138 Mus muscu	c 218	15.8	79.0	2614	6	BD057048	BD057048 A single
c 146	16.4	82.0	216763	2	AC145176	AC145176 Gorilla g	c 219	15.8	79.0	3037	10	AF183960	AF183960 Mus muscu
c 147	16.4	82.0	217860	10	AC123714	AC123714 Mus muscu	c 220	15.8	79.0	3071	9	BC023615	BC023615 Homo sapi
c 148	16.4	82.0	219181	2	AC141063	AC141063 Homo sapi	c 221	15.8	79.0	3384	9	BC064486	BC064486 Homo sapi
c 149	16.4	82.0	219359	10	AC110250	AC110250 Mus muscu	c 222	15.8	79.0	4384	9	BC080560	BC080560 Homo sapi
c 150	16.4	82.0	222536	2	AC117919	AC117919 Rattus no	c 223	15.8	79.0	4384	9	BC049838	BC049838 Homo sapi
c 151	16.4	82.0	225043	2	AC105657	AC105657 Rattus no	c 224	15.8	79.0	4766	10	RNGDNPN1	X71791 Rattus norv
c 152	16.4	82.0	225551	2	AC150560	AC150560 Mus muscu	c 225	15.8	79.0	5000	6	AR526834	AR526834 Sequence
c 153	16.4	82.0	227211	2	CR847561	CR847561 Danio rer	c 226	15.8	79.0	5174	10	BC060688	BC060688 Mus muscu
c 154	16.4	82.0	227904	2	AC139288	AC139288 Homo sapi	c 227	15.8	79.0	5334	6	CQ721623	CQ721623 Sequence
c 155	16.4	82.0	229277	2	AC127672	AC127672 Rattus no	c 228	15.8	79.0	5823	6	AX053263	AX053263 Sequence
c 156	16.4	82.0	229365	2	AC106518	AC106518 Rattus no	c 229	15.8	79.0	5823	9	D80007	D80007 Homo sapien
c 157	16.4	82.0	231295	2	AC097286	AC097286 Rattus no	c 230	15.8	79.0	5896	10	AL929468	AL929468 Mouse DNA
c 158	16.4	82.0	232538	2	AC114134	AC114134 Rattus no	c 231	15.8	79.0	7227	8	AY298951	AY298951 Chlamydom
c 159	16.4	82.0	234413	2	AC106066	AC106066 Rattus no	c 232	15.8	79.0	10801	10	AF215896	AF215896 Mus muscu
c 160	16.4	82.0	234651	2	AC115456	AC115456 Rattus no	c 233	15.8	79.0	11849	1	AE004500	AE004500 Pseudomon
c 161	16.4	82.0	236159	2	AC125553	AC125553 Rattus no	c 234	15.8	79.0	14289	1	AE007209	AE007209 Sinothizo
c 162	16.4	82.0	238624	2	AC112875	AC112875 Rattus no	c 235	15.8	79.0	15621	9	AC011652	AC011652 Homo sapi
c 163	16.4	82.0	238751	2	AC095116	AC095116 Rattus no	c 236	15.8	79.0	41511	9	AC131391	AC131391 Homo sapi
c 164	16.4	82.0	239394	2	AC110220	AC110220 Mus muscu	c 237	15.8	79.0	43658	14	HSV3PRGEN	M86409 Herpesvirus
c 165	16.4	82.0	240453	2	AC097603	AC097603 Rattus no	c 238	15.8	79.0	53130	2	AC101066	AC101066 Mus muscu

C 239	15.8	79.0	53795	6	AX695749	AX695749 Sequence	312	15.8	79.0	179390	10	AC122190	AC122190 Mus muscu
C 240	15.8	79.0	61004	2	BX927305_3	Continuation (4 of	313	15.8	79.0	179439	2	AC131897	AC131897 Oryctolag
C 241	15.8	79.0	66624	2	AC101520	Mus muscu	C 314	15.8	79.0	181054	2	AC149084	Mus muscu
C 242	15.8	79.0	73465	1	H8256611	AL109748 Homo sapi	C 315	15.8	79.0	181731	2	CR792458	CR792458 Danio rer
C 243	15.8	79.0	75131	2	AC023022	AL135799 Homo sapi	C 316	15.8	79.0	182650	2	AC141092	Rattus no
C 244	15.8	79.0	79389	2	AL135799	AL135799 Homo sapi	C 317	15.8	79.0	182809	10	AC135859	Mus muscu
C 245	15.8	79.0	94319	9	AC112501	AL12501 Homo sapi	C 318	15.8	79.0	183051	2	AC120130	Mus muscu
C 246	15.8	79.0	92458	9	AL353715	AL353715 Human DNA	C 319	15.8	79.0	183187	9	BS000003	BS000003 Pan trogl
C 247	15.8	79.0	93431	9	HU095737	U95737 Human Chrom	C 320	15.8	79.0	183462	10	AC129315	AC129315 Mus muscu
C 248	15.8	79.0	93519	9	AC130458	AL130458 Homo sapi	C 321	15.8	79.0	184164	10	AC126944	Mus muscu
C 249	15.8	79.0	103563	10	AL929313	AL929313 Mouse DNA	C 322	15.8	79.0	184194	9	AC011257	AC011257 Homo sapi
C 250	15.8	79.0	103826	9	AL139339	AL139339 Human DNA	C 323	15.8	79.0	184289	8	AP003545	AP003545 Oryza sat
C 251	15.8	79.0	106093	9	AC007561	AC007561 Homo sapi	C 324	15.8	79.0	185666	2	AC101747	AC101747 Mus muscu
C 252	15.8	79.0	106826	2	AC020829	AC020829 Mus muscu	C 325	15.8	79.0	185871	2	AC109251	AC109251 Mus muscu
C 253	15.8	79.0	110000	1	AE000516_04	Continuation (5 of	C 326	15.8	79.0	186120	9	AC007598	AC007598 Homo sapi
C 254	15.8	79.0	110000	8	CR382131_27	Continuation (28 o	C 327	15.8	79.0	186347	2	AC107710	Mus muscu
C 255	15.8	79.0	111934	9	AF281074	AF281074 Homo sapi	C 328	15.8	79.0	186431	10	AC113092	AC113092 Mus muscu
C 256	15.8	79.0	112544	9	AC109892	AC109892 Homo sapi	C 329	15.8	79.0	186517	9	AC022710	Homo sapi
C 257	15.8	79.0	112930	14	HSGEND	X64346 Salmiriline	C 330	15.8	79.0	190400	2	AC102726	AC102726 Mus muscu
C 258	15.8	79.0	113027	14	SHE410493	AU410493 Salmirilin	C 331	15.8	79.0	190421	10	AC132419	AC132419 Mus muscu
C 259	15.8	79.0	113626	2	AC145363	AC145363 Dasyopus n	C 332	15.8	79.0	190735	2	AC026084	AC026084 Homo sapi
C 260	15.8	79.0	114793	9	AF217796	AF217796 Homo sapi	C 333	15.8	79.0	191439	10	AL627302	AL627302 Mouse DNA
C 261	15.8	79.0	118999	6	AR526865	AR526865 Sequence	C 334	15.8	79.0	191809	10	AC117615	AC117615 Mus muscu
C 262	15.8	79.0	118999	9	AP240786	AP240786 Homo sapi	C 335	15.8	79.0	193927	2	AC120275	AC120275 Rattus no
C 263	15.8	79.0	118999	9	AP000351	AP000351 Homo sapi	C 336	15.8	79.0	194640	10	AL833775	AL833775 Mouse DNA
C 264	15.8	79.0	122461	10	AL645684	AL645684 Mouse DNA	C 337	15.8	79.0	194997	2	AC124670	AC124670 Mus muscu
C 265	15.8	79.0	135689	9	AC023755	AC023755 Homo sapi	C 338	15.8	79.0	197411	10	AC135961	AC135961 Mus muscu
C 266	15.8	79.0	135851	9	AC091714	AC091714 Papio anu	C 339	15.8	79.0	198294	10	AC135104	AC135104 Mus muscu
C 267	15.8	79.0	140819	10	AC121886	AC121886 Mus muscu	C 340	15.8	79.0	198992	10	AC109611	AC109611 Mus muscu
C 268	15.8	79.0	141809	10	AC125012	AC125012 Mus muscu	C 341	15.8	79.0	199140	2	BX005251	BX005251 Mus muscu
C 269	15.8	79.0	142172	2	AC015594	AC015594 Homo sapi	C 342	15.8	79.0	200229	10	AC099572	AC099572 Mus muscu
C 270	15.8	79.0	143420	27	AC102709	AC102709 Mus muscu	C 343	15.8	79.0	203057	10	AC099414	AC099414 Mus muscu
C 271	15.8	79.0	143074	2	AC102222	AC102222 Mus muscu	C 344	15.8	79.0	206592	10	AC124271	AC124271 Mus muscu
C 272	15.8	79.0	144037	2	RN484K6	AL603721 Rattus no	C 345	15.8	79.0	207275	2	AC130506	Rattus no
C 273	15.8	79.0	152244	9	AP000352	AP000352 Homo sapi	C 346	15.8	79.0	207312	10	AC110530	AC110530 Mus muscu
C 274	15.8	79.0	153170	9	HS1103G7	AL034548 Human DNA	C 347	15.8	79.0	207424	10	AC090489	AC090489 Genomic s
C 275	15.8	79.0	153907	9	AC146761	AC146761 Pan trogl	C 348	15.8	79.0	208207	10	AC145450	AC145450 Mus muscu
C 276	15.8	79.0	155041	2	CR450843	CR450843 Danio rer	C 349	15.8	79.0	212544	2	AC112790	AC112790 Mus muscu
C 277	15.8	79.0	157219	2	BX000348	BX000348 Danio rer	C 350	15.8	79.0	212835	2	AC007612	AC007612 Homo sapi
C 278	15.8	79.0	157757	10	AC131189	AC131189 Mus muscu	C 351	15.8	79.0	213687	2	AC119257	AC119257 Mus muscu
C 279	15.8	79.0	157860	9	AL391241	AL391241 Human DNA	C 352	15.8	79.0	213980	2	AC138547	AC138547 Mus muscu
C 280	15.8	79.0	159020	9	AL450307	AL450307 Human DNA	C 353	15.8	79.0	214031	2	AC134141	Rattus no
C 281	15.8	79.0	161119	8	AP004739	AP004739 Oryza sat	C 354	15.8	79.0	215983	2	AC141530	Rattus no
C 282	15.8	79.0	162090	10	AL805921	AL805921 Mouse DNA	C 355	15.8	79.0	216494	2	AC108336	Rattus no
C 283	15.8	79.0	164628	2	AC073646	AC073646 Homo sapi	C 356	15.8	79.0	216888	2	AC104918	Mus muscu
C 284	15.8	79.0	165033	2	AC087125	AC087125 Homo sapi	C 357	15.8	79.0	220434	2	AC141151	Rattus no
C 285	15.8	79.0	165332	2	AC011680	AC011680 Homo sapi	C 358	15.8	79.0	221859	10	AL732309	Mouse DNA
C 286	15.8	79.0	166604	10	AC125168	AC125168 Mus muscu	C 359	15.8	79.0	222605	2	AC095796	AC095796 Rattus no
C 287	15.8	79.0	167082	2	AC122555	AC122555 Mus muscu	C 360	15.8	79.0	222775	10	AC084069	AC084069 Mus muscu
C 288	15.8	79.0	167285	2	AC132853	AC132853 Mus muscu	C 361	15.8	79.0	223085	2	AC111419	Rattus no
C 289	15.8	79.0	167842	10	AC146603	AC146603 Mus muscu	C 362	15.8	79.0	223729	2	AC120905	Rattus no
C 290	15.8	79.0	168153	2	AC131519	AC131519 Rattus no	C 363	15.8	79.0	223729	2	AC120905	Rattus no
C 291	15.8	79.0	168337	2	AC118956	AC118956 Rattus no	C 364	15.8	79.0	223827	2	AC079441	AC079441 Mus muscu
C 292	15.8	79.0	169465	10	AL6065976	AL6065976 Mouse DNA	C 365	15.8	79.0	225687	2	AC098771	AC098771 Rattus no
C 293	15.8	79.0	169536	2	AC016496	AC016496 Homo sapi	C 366	15.8	79.0	226348	2	AC151108	AC151108 Mus muscu
C 294	15.8	79.0	169838	10	AL713861	AL713861 Mouse DNA	C 367	15.8	79.0	226714	2	AC079532	Mus muscu
C 295	15.8	79.0	170022	2	AC120072	AC120072 Rattus no	C 368	15.8	79.0	226736	10	AC098881	AC098881 Mus muscu
C 296	15.8	79.0	170323	10	AC124459	AC124459 Mus muscu	C 369	15.8	79.0	226809	2	AC109563	AC109563 Rattus no
C 297	15.8	79.0	170563	9	AC027020	AC027020 Homo sapi	C 370	15.8	79.0	229045	2	AC126138	Rattus no
C 298	15.8	79.0	170935	2	AC016079	AC016079 Homo sapi	C 371	15.8	79.0	230353	2	AC130642	Rattus no
C 299	15.8	79.0	171551	8	OSJN00177	AL662976 Oryza sat	C 372	15.8	79.0	231484	2	AC096379	AC096379 Rattus no
C 300	15.8	79.0	172380	10	AC126252	AC126252 Mus muscu	C 373	15.8	79.0	231767	2	AC094531	AC094531 Rattus no
C 301	15.8	79.0	172405	9	AC053523	AC053523 Homo sapi	C 374	15.8	79.0	232278	2	AC087876	AC087876 Mus muscu
C 302	15.8	79.0	174373	9	AC132216	AC132216 Homo sapi	C 375	15.8	79.0	232945	2	AC103156	AC103156 Rattus no
C 303	15.8	79.0	174833	2	AC127686	AC127686 Mus muscu	C 376	15.8	79.0	233785	2	AC105715	AC105715 Rattus no
C 304	15.8	79.0	174860	2	AC149560	AC149560 Papio anu	C 377	15.8	79.0	235198	2	AC106222	Rattus no
C 305	15.8	79.0	175727	10	AL732580	AL732580 Mus muscu	C 378	15.8	79.0	235405	10	AC101737	AC101737 Mus muscu
C 306	15.8	79.0	176027	2	AL732575	AL732575 Mus muscu	C 379	15.8	79.0	236011	2	AC105523	Rattus no
C 307	15.8	79.0	176384	10	AC114651	AC114651 Mus muscu	C 380	15.8	79.0	236635	2	AC126525	AC126525 Rattus no
C 308	15.8	79.0	178181	9	AC064868	AC064868 Homo sapi	C 381	15.8	79.0	236692	2	AC073773	AC073773 Mus muscu
C 309	15.8	79.0	178414	2	AC137801	AC137801 Homo sapi	C 382	15.8	79.0	238079	2	AC096000	AC096000 Rattus no
C 310	15.8	79.0	178957	2	AC151075	AC151075 Bos tauru	C 383	15.8	79.0	238776	2	AC126581	AC126581 Rattus no
C 311	15.8	79.0	179386	10	BX255910	BX255910 Mouse DNA	C 384	15.8	79.0	238877	2	AC115347	AC115347 Rattus no

C 385	15.8	79.0	242008	2	AC118601	AC118601 Mus muscu
C 386	15.8	79.0	244538	2	AC105696	AC105696 Rattus no
C 387	15.8	79.0	244567	2	AC118950	AC118950 Rattus no
C 388	15.8	79.0	245424	2	AC113069	AC113069 Rattus no
C 389	15.8	79.0	245977	2	AC115813	AC115813 Mus muscu
C 390	15.8	79.0	247112	2	AC097543	AC097543 Rattus no
C 391	15.8	79.0	247655	10	AC145549	AC145549 Mus muscu
C 392	15.8	79.0	248817	2	AC096230	AC096230 Rattus no
C 393	15.8	79.0	249291	2	AC113460	AC113460 Mus muscu
C 394	15.8	79.0	251133	2	AC121633	AC121633 Rattus no
C 395	15.8	79.0	253387	2	AC094775	AC094775 Rattus no
C 396	15.8	79.0	256717	2	AC123292	AC123292 Rattus no
C 397	15.8	79.0	256759	2	AC129832	AC129832 Rattus no
C 398	15.8	79.0	259889	2	AC125704	AC125704 Rattus no
C 399	15.8	79.0	261414	2	AC122960	AC122960 Rattus no
C 400	15.8	79.0	261968	2	AC120714	AC120714 Rattus no
C 401	15.8	79.0	262074	2	AC098624	AC098624 Rattus no
C 402	15.8	79.0	263505	10	AC139673	AC139673 Mus muscu
C 403	15.8	79.0	266517	2	AC107094	AC107094 Rattus no
C 404	15.8	79.0	267287	2	AC096702	AC096702 Rattus no
C 405	15.8	79.0	272166	2	AC120907	AC120907 Rattus no
C 406	15.8	79.0	273908	2	AL162851	AL162851 Homo sapi
C 407	15.8	79.0	275369	2	AC129357	AC129357 Rattus no
C 408	15.8	79.0	278583	10	AC132140	AC132140 Mus muscu
C 409	15.8	79.0	306172	2	AC103326	AC103326 Rattus no
C 410	15.8	79.0	324050	1	EX248335	EX248335 Mycobacte
C 411	15.8	79.0	324668	2	AC112334	AC112334 Rattus no
C 412	15.8	79.0	335598	2	AC133302	AC133302 Rattus no
C 413	15.8	79.0	340000	9	HS21C004	AL163204 Homo sapi
C 414	15.8	79.0	340493	10	EX883044	EX883044 Rattus no
C 415	15.8	79.0	342416	1	EX842573	EX842573 Mycobacte
C 416	15.4	77.0	266	9	HS318046	HS318046 Homo sapi
C 417	15.4	77.0	400	11	GI7953	GI7953 SHGC-14188
C 418	15.4	77.0	503	11	HSJ680C3	AL159883 STS from
C 419	15.4	77.0	792	9	HS3338372	AJ338372 Homo sapi
C 420	15.4	77.0	962	6	AX364857	AX364857 Sequence
C 421	15.4	77.0	1056	1	AY162311	AY162311 Unculture
C 422	15.4	77.0	1184	9	HS2299441	AJ299441 Homo sapi
C 423	15.4	77.0	1194	6	CQ721169	CQ721169 Sequence
C 424	15.4	77.0	1194	9	CR457354	CR457354 Homo sapi
C 425	15.4	77.0	1285	6	AX404151	AX404151 Sequence
C 426	15.4	77.0	1285	6	AX404153	AX404153 Sequence
C 427	15.4	77.0	2010	9	HS2299403	AJ299403 Homo sapi
C 428	15.4	77.0	2016	9	AK025998	AK025998 Homo sapi
C 429	15.4	77.0	2023	9	BC011746	BC011746 Homo sapi
C 430	15.4	77.0	2028	9	BC018292	BC018292 Homo sapi
C 431	15.4	77.0	2056	9	BC001085	BC001085 Homo sapi
C 432	15.4	77.0	2074	6	BD194882	BD194882 86 human
C 433	15.4	77.0	2074	6	CQ855210	CQ855210 Sequence
C 434	15.4	77.0	2081	9	BC007571	BC007571 Homo sapi
C 435	15.4	77.0	15244	9	HS344H3	Z83854 Human DNA s
C 436	15.4	77.0	19940	10	AB02208482	AB022085 Mus muscu
C 437	15.4	77.0	47461	9	AC141135	AC141135 Rattus no
C 438	15.4	77.0	64475	9	AF283320S2	AF283321 Homo sapi
C 439	15.4	77.0	68484	8	NCB17C10	AL355926 Neurospor
C 440	15.4	77.0	81368	2	AC094244_4	Continuation (5 of
C 441	15.4	77.0	83234	2	AC034104	AC034104 Homo sapi
C 442	15.4	77.0	107238	9	AC106786	AC106786 Homo sapi
C 443	15.4	77.0	109071	9	CR626912	CR626912 Human DNA
C 444	15.4	77.0	111363	2	AC008606	AC008606 Homo sapi
C 445	15.4	77.0	115036	2	AP000444	AP000444 Homo sapi
C 446	15.4	77.0	115141	8	AC145324	AC145324 Oryza sat
C 447	15.4	77.0	116730	9	AC008548	AC008548 Homo sapi
C 448	15.4	77.0	119916	9	AF241734	AF241734 Homo sapi
C 449	15.4	77.0	120166	9	AC091390	AC091390 Homo sapi
C 450	15.4	77.0	125209	10	AC079871	AC079871 Mus muscu
C 451	15.4	77.0	128910	10	AC125308	AC125308 Mus muscu
C 452	15.4	77.0	129282	2	AC148573	AC148573 Taenlopyg
C 453	15.4	77.0	135431	9	AC106847	AC106847 Homo sapi
C 454	15.4	77.0	136155	10	AC132586	AC132586 Mus muscu
C 455	15.4	77.0	137206	2	AC024123	AC024123 Homo sapi
C 456	15.4	77.0	137981	10	AC135108	AC135108 Mus muscu
C 457	15.4	77.0	142863	9	AL512631	AL512631 Human DNA
C 458	15.4	77.0	143241	10	CNS07Y0X	AL713881 Mus muscu
C 459	15.4	77.0	147031	2	AC144505	AC144505 Pan trogl
C 460	15.4	77.0	148657	2	AC116809	AC116809 Mus muscu
C 461	15.4	77.0	149428	2	AC010264	AC010264 Homo sapi
C 462	15.4	77.0	150507	9	AC107918	AC107918 Homo sapi
C 463	15.4	77.0	157756	2	AP001075	AP001075 Homo sapi
C 464	15.4	77.0	157866	2	AL139132	AL139132 Human DNA
C 465	15.4	77.0	158420	9	AL732326	AL732326 Human DNA
C 466	15.4	77.0	162685	5	EX248131	EX248131 Zebrafish
C 467	15.4	77.0	163652	2	AC145845	AC145845 Pan trogl
C 468	15.4	77.0	164055	2	AC130883	AC130883 Rattus no
C 469	15.4	77.0	167123	10	AL772277	AL772277 Mouse DNA
C 470	15.4	77.0	167284	9	AL112719	AL112719 Homo sapi
C 471	15.4	77.0	168872	2	AC145844	AC145844 Pan trogl
C 472	15.4	77.0	170831	9	AC009055	AC009055 Homo sapi
C 473	15.4	77.0	172420	9	AC022182	AC022182 Homo sapi
C 474	15.4	77.0	174425	2	AC146284	AC146284 Lemur cat
C 475	15.4	77.0	175706	2	AC013798	AC013798 Homo sapi
C 476	15.4	77.0	175749	9	AC025857	AC025857 Homo sapi
C 477	15.4	77.0	176191	2	AC131060	AC131060 Mus muscu
C 478	15.4	77.0	176641	2	AC084869	AC084869 Homo sapi
C 479	15.4	77.0	177694	2	AC019358	AC019358 Homo sapi
C 480	15.4	77.0	177728	9	AL162425	AL162425 Human DNA
C 481	15.4	77.0	178023	2	AC115945	AC115945 Mus muscu
C 482	15.4	77.0	178364	10	AL645594	AL645594 Mouse DNA
C 483	15.4	77.0	178607	9	AC007068	AC007068 Homo sapi
C 484	15.4	77.0	180541	2	AC016521	AC016521 Homo sapi
C 485	15.4	77.0	182770	9	AC104082	AC104082 Homo sapi
C 486	15.4	77.0	184450	2	AC116417	AC116417 Mus muscu
C 487	15.4	77.0	186626	9	AP000902	AP000902 Homo sapi
C 488	15.4	77.0	186780	9	AC005740	AC005740 Homo sapi
C 489	15.4	77.0	187101	9	AL589862	AL589862 Human DNA
C 490	15.4	77.0	187415	2	AC110802	AC110802 Homo sapi
C 491	15.4	77.0	189550	10	AC140465	AC140465 Mus muscu
C 492	15.4	77.0	191588	2	AC132030	AC132030 Rattus no
C 493	15.4	77.0	195514	2	AC110737	AC110737 Mus muscu
C 494	15.4	77.0	195972	10	AC117567	AC117567 Mus muscu
C 495	15.4	77.0	198720	9	AP002366	AP002366 Homo sapi
C 496	15.4	77.0	199275	9	AC011401	AC011401 Homo sapi
C 497	15.4	77.0	199665	9	AF064857	AF064857 Homo sapi
C 498	15.4	77.0	199719	2	AC127611	AC127611 Rattus no
C 499	15.4	77.0	200735	2	AC110330	AC110330 Rattus no
C 500	15.4	77.0	203768	2	AC146069	AC146069 Pan trogl
RESULT 1						
CQ731079/c						
LOCUS	CQ731079	1581 bp	DNA	linear	PAT 03-FEB-2004	
DEFINITION	Sequence 17013 from Patent WO02068579.					
ACCESSION	CQ731079					
VERSION	CQ731079.1	GI:42306695				
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.					
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof					
JOURNAL	Patent: WO 02068579-A 17013 06-SEP-2002; PE Corporation (NY) (US)					
FEATURES	Location/Qualifiers					
source	1..1581					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						



```

Query Match      100.0%; Score 20; DB 6; Length 1381;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
|||||
Db 949 AGGGCGTCTCTGAGTAGCAG 930

RESULT 2
LOCUS HSU59269 2088 bp mRNA linear PRI 24-SEP-1996
DEFINITION Human hyaluronan synthase mRNA, complete cds.
ACCESSION U59269
VERSION U59269.1 GI:1556464
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2088)
AUTHORS Shyan,A.M., Heldin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
TITLE Functional cloning of the cDNA for a human hyaluronan synthase
JOURNAL J. Biol. Chem. 271 (38), 23395-23399 (1996)
MEDLINE 96394438
PUBMED 8798544
REFERENCE 2 (bases 1 to 2088)
AUTHORS Briskin,M.J. and Shyan,A.M.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
Cambridge, MA 02142, USA
FEATURES
source
1..2088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
36..1772
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAC50706.1"
/db_xref="GI:1556465"
/translation="MRQDAKPTPAARRCSGLARRVUTIAFALLILGLMTWYAAAGV
PLASDRYGLAFGLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDAAATARSVALTIS
AYQEDPAYLRQCLASARALYPRARLVLMVVDGNRAEDLYMDMFREVFADDPATY
VWDGNYHPWPAAGAVGAGAREVEADPGRLAVEALVTRRCVCVAORWGKREY
MYTAFKALGDSVDVYQVCDSDTRLPDMLLELVRLDEDPVAGVGDDVRLINPLDSM
VSFSLSLRYWAFNVERACQSYFHCVSCISGPLGLYRNLLQQLFLEAWYKQFLGTHC
TFDGRHLTNRLSMGYATKYTSRSCYSETPSSFLRWLSQQTRWSKYFREWLYNAL
WHRHRAWMTYEAVVSGLPFPFVAATVLRFLYAGRPWALLWLLCVQVGLAKAAFAFA
WLRGLRMVLLSLYAPLYMCGLLPAKFLALVTMNSQSGWGTSGRRKLAANYVPLPLAL
WALLLLGLVRSVAHEARADWSGPRAAEAYHLAAGAGAYGVYVWVAMLTLYWVGVRRL
CRRRTGGYRVQV"
ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 2088;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
|||||
Db 1134 AGGGCGTCTCTGAGTAGCAG 1115

RESULT 3
HUMHAS/c
LOCUS HUMHAS 2108 bp mRNA linear PRI 06-FEB-1999
DEFINITION Homo sapiens mRNA for hyaluronan synthase, complete cds.
ACCESSION D84424
VERSION D84424.1 GI:1401033
KEYWORDS hHAS; hyaluronan synthase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2108)
AUTHORS Itano,N. and Kimata,K.
TITLE Molecular cloning of human hyaluronan synthase
JOURNAL Biochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)
MEDLINE 96244584
PUBMED 8651928
REFERENCE 2 (bases 1 to 2108)
AUTHORS Itano,N.
JOURNAL Unpublished
AUTHORS Itano,N.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine; Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-364-4811(ex.2087), Fax:0561-63-3532)
FEATURES
source
1..2108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="brain"
/clone_lib="/lgtl1"
/dev_stage="fetal"
1..2108
/gene="hHAS"
149..1780
/gene="hHAS"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="BAAL2351.1"
/db_xref="GI:1401034"
/translation="MTWYAAAGVPLASDRYGLAFGLYGAFLSAHLVAQSLFAYLEHHR
RVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARALYPRARVRLMVVDGN
RAEDLYMDMFREVFADDPATYVWDGNYHPWPAAGAVGAGAREVEADPGRLA
VEALVTRRCVCVAORWGKREYMTAFKALGDSVDVYQVCDSDTRLPDMLLELVRL
LDEDPRVAGVGDDVRLINPLDSWVSFLSLRYWAFNVERACQSYFHCVSCISGPLGL
YRNLLQQLFLEAWYKQFLGTHCTFGDGRHLTNRLSMGYATKYTSRSCYSETPSSFL
LRWLSQQTRWSKYFREWLYNALWHRHRAWMTYEAVVSGLPFPFVAATVLRFLYAGR
PWALLWLLCVQVGLAKAAFAAALRGCLRVLLSLYAPLYMCGLLPAKFLALVTMNSQ
SGWGTSGRRKLAANYVPLPLALWALLLLGLVRSVAHEARADWSGPRAAEAYHLAAG
GAGAYGVYVWVAMLTLYWVGVRRLCRRRTGGYRVQV"
ORIGIN
Query Match      100.0%; Score 20; DB 9; Length 2108;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
|||||
Db 1142 AGGGCGTCTCTGAGTAGCAG 1123

RESULT 4
LOCUS AR220003/c
DEFINITION Sequence 1 from patent US 6423514.
ACCESSION AR220003
VERSION AR220003.1 GI:23324403
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2116)
AUTHORS Briskin,M.J.
TITLE Mammalian hyaluronan synthases, nucleic acids and uses thereof
JOURNAL Patent: US 6423514-A 1 23-JUL-2002;
FEATURES
source
1..2116
/organism="unknown"
/mol_type="mRNA"

```

Query Match	100.0%;	Score 20;	DB 6;	Length 2116;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches	20;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1 AGGGCGTCTCTGAGTAGCAG 20			
Db	1134 AGGGCGTCTCTGAGTAGCAG 1115			
RESULT 5				
LOCUS	AR137038/c			
DEFINITION	Sequence 1 from patent US 6162908.			
ACCESSION	AR137038			
VERSION	AR137038.1 GI:14478288			
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2117)			
AUTHORS	Itano,N. and Kimata,K.			
TITLE	Polypeptide of human-origin hyaluronate synthetase and DNA encoding the same			
JOURNAL	Patent: US 6162908-A 1 19-DEC-2000;			
FEATURES	Location/Qualifiers			
source	1..2117			
	/organism="unknown"			
	/mol_type="unassigned DNA"			
ORIGIN				
Query Match	100.0%;	Score 20;	DB 6;	Length 2117;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches	20;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1 AGGGCGTCTCTGAGTAGCAG 20			
Db	1142 AGGGCGTCTCTGAGTAGCAG 1123			
RESULT 6				
LOCUS	BC035837/c			
DEFINITION	Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218 IMAGE:5889083), complete cds.			
ACCESSION	BC035837			
VERSION	BC035837.1 GI:23243101			
KEYWORDS	MGC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 2119)			
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,E., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalusz,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			

12477932	PUBMED			
2 (bases 1 to 2119)	REFERENCE			
Strausberg,R.	AUTHORS			
Direct Submission	TITLE			
Submitted (31-JUL-2002)	JOURNAL			
National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
NIH-MGC Project URL: http://mgc.nci.nih.gov	REMARK			
Contact: MGC help desk	COMMENT			
Email: cgapbs-r@mail.nih.gov				
Tissue Procurement: Invitrogen				
CDNA Library Preparation: Life Technologies, Inc.				
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)				
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;				
Web site: http://www.nisc.nih.gov/				
Contact: nisc_mgc@hgrl.nih.gov				
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsourgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov				
Series: IRAC Plate: 79 Row: F Column: 24				
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504338.				
Location/Qualifiers	FEATURES			
1..2119	source			
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="MGC:46218 IMAGE:5889083"				
/tissue_type="Ovary, pooled from 3 adults"				
/clone_lib="NIH_MGC_125"				
/lab_host="DH10B"				
/note="vector: pCMV-SPORT6"				
1..2119	gene			
/gene="HAS1"				
/note="synonym: HAS"				
/db_xref="LocusID:3036"				
/db_xref="MIM:601463"				
42..1775	CDS			
/gene="HAS1"				
/codon_start=1				
/product="hyaluronan synthase 1"				
/protein_id="AAH35837.1"				
/db_xref="GI:23243102"				
/db_xref="LocusID:3036"				
/db_xref="MIM:601463"				
/translation="MRQDAKPTPAARRCCGLARRVLTIAPALLILGLMTWAAGVP LASRYGLAFLGLYFELSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISA YQEDPAYLRQCLASARALLPRAKRLVLMVDNGRAEDLYWDMFEVFADEDPATVY WDTGNHQPWEPADYAAVGAGAYREVEADEPGLRAVEALVTRRCVCVAQRWGKREVM YTAFLKQDPMALDELVDLDEPRVGAVGGDVRILNPLDSMW SFLSRLRYAFNVERACQSYFHCVSCISGFLGLYRNRLIQQFLEAWYNKQFLGHT FGDDRHLTNRMLSMGYATKYTSRCYSCTPSSFLRWLSQOTRWSKSYFRELWYLNAL WHRHAWMTVEAVVUGLFPFPVATVLRLFVAGRPWALLVLLCVQGVALAKAAFAW LRGRIRWLLSLYAPLYTWCGLLPKFAFLVTMQSGWGTSGRKLKLANYPPLPLALW ALLLLGGLVRSVAHEARADWSGFSRAAEAYHLAAGAGAYVGYVWVAMLTLLYVWGVRRIC RRTTGGYRVQV"				
ORIGIN				
Query Match	100.0%;	Score 20;	DB 9;	Length 2119;
Best Local Similarity	100.0%;	Pred. No. 44;		
Matches	20;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			

```

Qy 1 AGGGCGTCTCTGAGTAGCAG 20
Db 1137 AGGGCGTCTCTGAGTAGCAG 1118

RESULT 7
AC130783 193986 bp DNA linear HTG 19-NOV-2002
LOCUS Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8
DEFINITION ordered pieces.
ACCESSION AC130783
VERSION AC130783.2 GI:25100968
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
REFERENCE 1 (bases 1 to 193986)
AUTHORS Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakeley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Marques,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Duque,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Weherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
UNPUBLISHED
JOURNAL NISC Comparative Sequencing Initiative
REFERENCE 2 (bases 1 to 193986)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 193986)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
COMMENT On Nov 19, 2002 this sequence version replaced gi:22218453.
----- Genomc Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dxd
Center clone name: 426A12

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 191764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11512: contig of 11512 bp in length
* 11513 11612: gap of unknown length
* 11613 59488: contig of 47876 bp in length
* 59489 59588: gap of unknown length
* 59589 103266: contig of 43678 bp in length
* 103267 103366: gap of unknown length
* 103367 127885: contig of 24519 bp in length
* 127886 127986: gap of unknown length
* 127987 129692: contig of 1707 bp in length
* 129693 129792: gap of unknown length
* 129793 131077: contig of 1285 bp in length
* 131078 131177: gap of unknown length
* 131178 174663: contig of 43486 bp in length
* 174664 174763: gap of unknown length
* 174764 193986: contig of 19223 bp in length.
FEATURES
Location/Qualifiers
1. 193986
/organism="pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-426A12"
/clone_lib="CH251"
misc_feature 1. 11512
/feature="assembly_fragment"
clone_end:SP6
vector_side:left
1. 76
/feature="assembly_fragment"
misc_feature 11613..59488
/feature="assembly_fragment"
misc_feature 59589..103266
/feature="assembly_fragment"
misc_feature 103367..127885
/feature="assembly_fragment"
misc_feature 127986..129692
/feature="assembly_fragment"
misc_feature 129793..131077
/feature="assembly_fragment"
misc_feature 131178..174663
/feature="assembly_fragment"
misc_feature 174764..193986
/feature="assembly_fragment"
clone_end:T7
vector_side:right
ORIGIN
Query Match 100.0%; Score 20; DB 2; Length 193986;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGCGTCTCTGAGTAGCAG 20
Db 128496 AGGGCGTCTCTGAGTAGCAG 128515

RESULT 8
AC018755
LOCUS AC018755 Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
DEFINITION sequence.
ACCESSION AC018755
VERSION AC018755.3 GI:9454515

```

```

KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 229155)
AUTHORS Lamerdin, J.E., McCreedy, P.M., Skowronski, E., Viswanathan, V.,
Burkhart-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stilwagen, S.,
Phan, H., Velasco, N., Do, L., Regala, N., Terry, A., Brower, A.,
Garnes, J., Danganan, L., Erler, A., Christensen, M., Georgescu, A.,
Avila, J., Liu, S., Andreise, J., Trankheim, M., Attix, C.,
Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R.,
Thomas, P., Quan, G., Krommiller, B., Arellano, A., Sanders, C., Ow, D.,
Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V.
Sequence analysis of a 5-Mb region in 19q13.4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 229155)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 229155)
AUTHORS Lamerdin, J.E.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence
Livermore National Laboratory, Livermore, CA 94550
COMMENT On Jul 26, 2000 this sequence version replaced gi:7458780.
Map and sequence oriented from centromere to q-telomere. BC330783
(CIT-HSPC 470E3) is overlapped on the left by BC849408
(CITB-E1 3073N11, AC020914) from bases 1 to 125 of this accession,
and overlaps cosmid R28782 (LLNL-R 248F10, AC005946) on the right
from bases 217,905 to 229,155 of this accession. Additional
chromosome 19 map and sequence information may be obtained at:
http://www-bio.llnl.gov/bbrp/genome/genome.html.
FEATURES
Source
1..229155
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="19"
/map="19q13.4"
/clone="CITC-470E3 (LLNL clone name BC330783)"
/clone_lib="Caltech human BAC library C"
repeat_region 201..262
/rpt_family="L2"
gene <299..16870
/genes="ZNF175"
/notes="zinc finger protein 175; synonym: OTK18"
join(<299..550,8540..8666,8936..9031,13776..16870)
/genes="ZNF175"
/product="ZNF175 mRNA"
join(479..550,8540..8666,8936..9031,13776..15616)
/genes="ZNF175"
/codon_start=1
/product="zinc finger protein 175"
/db_xref="GI:9454517"
/protein_id="AAF87843.1"
translation="MPADVNLGKQPVGLPEKQDGCSEASVFEDVTVDPSREWQQL
DPAQRCLYRDVLMLEYLSHVFAGVYHYPNPEVIFMLKEKEPRVEAEVSHQRCQREF
GLEIPKETSXKASFOKVMGFTRDGSWCSILELRDLADRTKKDEQNIQPMGHSA
FNFKTLNTSECEYKDPCKMIRTRPHLASSOKPOKCLFTESLKLNEVNGNSEN
DTEQLDDVVGSGQLFSHSSDACSKNIHTGETCKGNQCKVCQGHQSKLQKHITQK
KPDGCEGQSFQSKSHLFAQQRHISVGNLHCGKCGKAPFQMLSLVLTHTDGP
CICKEGCVFIQRSHLLTHQTRKRYKCHDCGKAFQMLSLFRHQRTHREKLYE
CEGQAFQKALHIVHQRHTGEPYCHNCGKSFPSKSLDIHRIHTGEPKVCYVCI
BGTFTOKSHLHOKHTGERHHVCEGCAFNKSKILSMHORHTGEPKVCSECGK
ATTSKQFKEHRIHTGEPYCVTEGCKAFNGSRNPKHQIHTHTRERPVCYKCGKAP
VQSKSLHTHRTMGKPECLDCGKSFSGKQKLVKQRIHTGERPVCSECGKAPFN
RSNFKHQHTTRDYSKCSYKVGTKQ"
repeat_region 1097..1198
/rpt_family="MER99"
repeat_region 1559..1829

```

```

repeat_region /rpt_family="AluJb"
complement(2130..2355)
/rpt_family="L1ME"
repeat_region complement(2359..2627)
/rpt_family="AluY"
repeat_region complement(2816..3183)
/rpt_family="HAL1"
repeat_region complement(3290..3598)
/rpt_family="AluSg"
repeat_region complement(3611..3760)
/rpt_family="MIR"
repeat_region complement(3766..4067)
/rpt_family="HAL1"
repeat_region 4098..4215
/rpt_family="AluJb"
repeat_region complement(4339..4406)
/rpt_family="HAL1"
repeat_region complement(4518..4724)
/rpt_family="HAL1"
repeat_region complement(4803..5110)
/rpt_family="AluY"
repeat_region 5266..5561
/rpt_family="AluSg"
repeat_region 5635..5932
/rpt_family="AluJo"
repeat_region 5984..6120
/rpt_family="AluJb"
repeat_region 6121..6423
/rpt_family="AluSx"
repeat_region 6424..6579
/rpt_family="AluJb"
repeat_region 6647..6913
/rpt_family="L1ME"
repeat_region 6936..7295
/rpt_family="MER33"
repeat_region 7700..7785
/rpt_family="L2"
repeat_region complement(9216..9434)
/rpt_family="L1ME"
repeat_region complement(9435..9745)
/rpt_family="AluY"
repeat_region complement(9746..9815)
/rpt_family="L1ME"
repeat_region complement(10009..10303)
/rpt_family="AluSp"
repeat_region 11783..12637
/rpt_family="L1MB2"
STS 12180..12330
/gene="ZNF175"
repeat_region 13436..13619
/rpt_family="MIR"
repeat_region 16068..16112
/rpt_family="(TA)n"
repeat_region complement(16113..16329)
/rpt_family="MER46A"
repeat_region 16330..16390
/rpt_family="(TA)n"
repeat_region 16403..16439
/rpt_family="(TATG)n"
repeat_region 16793..17013
/rpt_family="MER58A"
repeat_region 17021..17052
/rpt_family="(CA)n"
repeat_region 17311..17545
/rpt_family="MIR"
repeat_region complement(18683..18896)
/rpt_family="MIR"
repeat_region complement(18915..19183)
/rpt_family="AluY"
repeat_region 19336..19599
/rpt_family="MIR"
repeat_region 19811..19928

```



strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/OsNj, BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J, MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei."  
<1..>600

# STSS

## ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 600;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
|||||  
Db 160 AGGGCGTCTCCGAGTAGCA 178  
|||||

# RESULT 11

## LOCUS

BV160504 600 bp DNA linear STS 15-MAY-2004  
RPMMS00038150 Roche Palo Alto Mus musculus STS genomic, sequence tagged site.

ACCESSION BV160504 GI:47263404  
VERSION BV160504  
KEYWORDS STS.

SOURCE Mus musculus (house mouse)

# ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 600)  
Usuka, J., Liao, G., Cheng, J., Nguyen, A., Bach, C., Puech, A., McPherson, J. D., Foerzler, D. and Peltz, G.  
Mus musculus SNPs  
Unpublished (2003)

# TITLE

## JOURNAL

### COMMENT

Contact: Jonathan Usuka  
Roche Palo Alto Genetics and Genomics Department  
Roche Palo Alto  
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA  
Tel: 6508555807  
Email: Jonathan.Usuka@roche.com  
Primer A: No primer submitted with this STS  
Primer B: No primer submitted with this STS.

# FEATURES

## source

1..600  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/map="17-7953-8459-CAAA01219574.1.1.20118"  
/clone\_lib="Roche Palo Alto"  
/note="SNPs developed from assay sequences derived from 15 different strains of mice (as of October 1, 2003). Those strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/OsNj, BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J, MRL/MpJ, NZB/BinJ, NZW/Lac, SPRET/Ei."  
<1..>600

# STSS

## ORIGIN

Query Match 87.0%; Score 17.4; DB 11; Length 600;  
Best Local Similarity 94.7%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
|||||  
Db 160 AGGGCGTCTCCGAGTAGCA 178  
|||||

# RESULT 12

## AB097568/c

### LOCUS

AB097568 2095 bp mRNA linear ROD 28-APR-2004  
Rattus norvegicus HAS1 mRNA for hyaluronan synthase 1, complete cds.

# ACCESSION

## AB097568

### VERSION

#### KEYWORDS

AB097568.1 GI:26453348

# SOURCE

## ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

# REFERENCE

## AUTHORS

Itano, N., Sawai, T., Atsumi, F., Miyaishi, O., Taniguchi, S.,

Kannagi, R., Hamauchi, M. and Kimata, K.

Selective expression and functional characteristics of three

Mammalian hyaluronan synthases in oncogenic malignant

transformation

J. Biol. Chem. 279 (18), 18679-18687 (2004)

14724275

# REFERENCE

## AUTHORS

### TITLE

#### JOURNAL

Itano, N.

Direct Submission

Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University,

Institute for Molecular Science of Medicine; Nagakute, Aichi, Aichi

480-1195, Japan [E-mail: itano@amugw.aichi-med-u.ac.jp,

Tel: 81-52-264-4811 (ex. 2095), Fax: 81-561-63-3532)

# FEATURES

## source

1..2095  
/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

# gene

## CDS

1..2095  
/gene="HAS1"  
43..11794  
/gene="HAS1"  
/codon\_start=1  
/product="hyaluronan synthase 1"  
/protein\_id="BAC43730.1"  
/db\_xref="GI:26453349"  
/translation="MRQMPKPESEAAACSCGLARRVLITTFALLILGLMTWAYAGVP  
LASDPYGLLAGLYGAFLSAHILYAOQLSFAYLEHRRVTVAARFAKGLDAAATARSVA  
LITISAYQDPTYLROCLTSARALLYPRTLRLVLMVVDGNRAEDLYMVDNREVFADDED  
PATYVMDGNVHQPWEPAPAAAGVAGGAYREVEADPEGLRALEALVTRRCVCVQARWG  
GKREVMYAFKALGDSVDYVOVCDSDTRLDPMALLELVRVLDEPRVGAAGDVRILN  
PLDSWVSFLSLRLYVAVNVERACQSYFHCUSCISGPIGLYRNLLLOQFLEAWYNOKF  
LGTHCTGDHRLTNRMLSMGYATIKYTSRSCISPPSFLRWLSQOTRWSKSYFREW  
LYNALWHRHHAWMTYEAVVGLFPFFVAATVLRFLYAGRPWALLWLVLCVQVALAK  
AAFAALRGCLRMVLLSLYAPLYMCGLLPAKFLALVTMNSQMGSGRKKLAANYVYV  
LPALWALLLGLLRSVAQVEVRADWSGSPRAAEAYHLAAGASAYVAYVIMLTIYV  
GVRLCRRRSGGYRVQV"

# ORIGIN

## Query Match

Best Local Similarity 94.7%; Pred. No. 8.2e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
|||||

Db 1156 AGGGCGTCTCTGAGTAGCA 1138  
|||||

# RESULT 13

## E13681/c

### LOCUS

E13681 DNA encoding novel mouse hyaluronate synthetase.

# DEFINITION

## ACCESSION

### VERSION

#### KEYWORDS

##### SOURCE

###### ORGANISM

###### REFERENCE

###### AUTHORS

###### TITLE

# JOURNAL

## COMMENT

E13681 2102 bp DNA linear PAT 27-APR-1998  
DNA encoding novel mouse hyaluronate synthetase.  
E13681  
E13681.1 GI:3252450  
JP 1997224674-A/1.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2102)  
Itano, N. and Kimata, H.  
POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING THE SAME  
Patent: JP 1997224674-A 1 02-SEP-1997;  
SCIENCE & TECH AGENCY  
OS Mus musculus (mouse)

```

PN JP 1997224674-A/1
PD 02-SEP-1997
PF 26-FEB-1996 JP 1996038336
PI ITANO NAOKI, KIMATA HIROHARU
PC C12N15/09, C07K14/47, C12N9/00//A61K48/00, C12N1/21, PC
(C12N9/00,
PC C12R1.91), (C12N9/00, C12R1.19), (C12N1/21, C12R1.19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2102
/organism="Mus musculus"
FT 5'UTR 1..48
/cell_line="FM3A"
FT CDS 49..1800
/feature="Novel mouse hyaluronate synthetase"
FT 3'UTR 1801..2102.
Location/Qualifiers
FT 1..2102
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 2102;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGAGTAGCA 19
Db 1162 AGGGCGTCTCCGAGTAGCA 1144
RESULT 14
LOCUS E28454/c 2102 bp DNA linear PAT 18-JUN-2001
DEFINITION Hyaluronate synthase promoter DNA.
ACCESSION E28454
VERSION E28454.1 GI:13018346
KEYWORDS JP 1999196875-A/2.
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Yoichi, Y., Naoki, I. and Koji, K.
JOURNAL Hyaluronate synthase promoter DNA
Patent: JP 1999196875-A 2 27-JUL-1999;
SEIKAGAKU KOGYO CO LTD
COMMENT OS Mus sp. (mouse)
PN JP 1999196875-A/2
PD 27-JUL-1999
PF 14-JAN-1998 JP 1998006191
PR YOICHI YAMADA, NAOKI ITANO, KOJI KIMATA
PC C12N15/09, C12N9/00, C12Q1/68// (C12N15/09, C12R1.91), C12N15/00,
(C12N15/00, C12R1.91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 49..1800.
Location/Qualifiers
FT 1..2102
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 2102;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGAGTAGCA 19
Db 1162 AGGGCGTCTCCGAGTAGCA 1144

```

```

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGAGTAGCA 19
Db 1162 AGGGCGTCTCCGAGTAGCA 1144
RESULT 15
LOCUS E30971/c 2102 bp DNA linear PAT 18-JUN-2001
DEFINITION Hyaluronate synthetase modified protein.
ACCESSION E30971
VERSION E30971.1 GI:13017286
KEYWORDS JP 2000004886-A/1.
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Naoki, I., Mamoru, Y. and Koji, K.
JOURNAL Hyaluronate synthetase modified protein
Patent: JP 2000004886-A 1 11-JAN-2000;
SEIKAGAKU KOGYO CO LTD
COMMENT OS Mus sp. (mouse)
PN JP 2000004886-A/1
PD 11-JAN-2000
PF 24-JUN-1998 JP 1998193788
PR NAOKI ITANO, MAMORU YOSHIDA, KOJI KIMATA
PC C12N15/09, C12N9/00, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 49..1800.
Location/Qualifiers
FT 1..2102
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 2102;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGAGTAGCA 19
Db 1162 AGGGCGTCTCCGAGTAGCA 1144
RESULT 16
LOCUS E34326/c 2102 bp DNA linear PAT 31-JAN-2002
DEFINITION DNA for gene targeting hyaluronate synthase gene.
ACCESSION E34326
VERSION E34326.1 GI:18624311
KEYWORDS JP 2000116382-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Itano, N. and Kimata, K.
JOURNAL DNA for gene targeting hyaluronate synthase gene
Patent: JP 2000116382-A 1 25-APR-2000;
SEIKAGAKU KOGYO CO LTD
COMMENT OS Mus musculus (mouse)
PN JP 2000116382-A/1
PD 25-APR-2000
PF 13-OCT-1998 JP 1998291201
PR NAOKI ITANO, KOJI KIMATA
PC C12N15/09, C12N5/10, G01N33/50//C12N9/00, C12Q1/68, (C12N15/09, PC

```

```

C12R1:91),
PC (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91),
PC (C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..2102
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

FEATURES
source
ORIGIN
Query Match 87.0%; Score 17.4; DB 6; Length 2102;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
|||||
Db 1162 AGGGCGTCTCCGAGTAGCA 1144

RESULT 17
MUSHAS/c
LOCUS MUSHAS 2102 bp mRNA linear ROD 06-FEB-1999
DEFINITION Mus musculus mRNA for hyaluronan synthase, complete cds.
ACCESSION D82964.1 GI:1339939
VERSION D82964.1
KEYWORDS hyaluronan synthase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (sites)
Itano,N. and Kimata,K.
Expression cloning and molecular characterization of HAS protein, a
eukaryotic hyaluronan synthase
J. Biol. Chem. 271 (17), 9875-9878 (1996)
MEDLINE 96215261
PUBMED 8626618
REFERENCE 2 (bases 1 to 2102)
AUTHORS Itano,N.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2102)
AUTHORS Itano,N.
JOURNAL Direct Submission
TITLE Submitted (08-JAN-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine; Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)
JOURNAL
FEATURES
source
Location/Qualifiers
1..2102
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="FM3A H41"
/clone_lib="cDNA/pCDNA1"
49..1800
/codon_start=1
/product="hyaluronan synthase"
/protein_id="BAA11654.1"
/db_xref="GI:1339940"
/translation="MRQDMKPSEAARCCSGLARRALTIIFALLILGLMTWAYAGVP
LASDRYGLLAFGLYGAFLSAHLAQSLPAYLEHRRVAAARSLAKGPDLAATARSVA
LTISAYQEDPAYLRQCLVSARALLYPHRLRLVMVDGNRPEDLYMVMDFREVFADDE
PATYVWDGNYPHPWEPAEATGAGVEGAYREVEADPGRLAVALVTRRCVCVQARWG
GKREVMYAFKALGSDVDYQVCDSDTRLDPMALLELVQLDEDPVGVAGGVDVRLIN
PLDSWVSFLSLRYWFAFNVACOSYFHCVSCI SGPLGLYRNLLQQFLKAWYNQKF
LGTHCTFGDDRHLTNRMLSMGYATKTSRSRCYSETPSSFLRWLSQOTRWSKSYFREW
LYNALWHRHHAWMTYEAVVSGLFFEFVAATVIRLFYAGRPWALLVLLVCVQVALAK
AAFAAWLRGCVRMVLLSYAPLYMCGLLPAKFLALVTMNSGNGTSGRKKLAANYVPV
LPALWALLLLGLLARSVAQEARADWSGFSRAAEAYHLLAAGAGAYYVYVWMLTIYVY
GVRRLCRRRSGGYRVQV"

CDS
ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 2109;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
|||||
Db 1149 AGGGTCTCTCGAGTAGCA 1131

RESULT 19
AC108651_3

```

```

Query Match 87.0%; Score 17.4; DB 10; Length 2102;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
|||||
Db 1162 AGGGCGTCTCCGAGTAGCA 1144

RESULT 18
AY463695/c
LOCUS AY463695 2109 bp mRNA linear PRI 07-DEC-2003
DEFINITION Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.
ACCESSION AY463695
VERSION AY463695.1 GI:38607341
KEYWORDS
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Papio.
REFERENCE
1 (bases 1 to 2109)
Martinez-Duncker,I., Oriol,R. and Mollicone,R.
Evolution of the the hyaluronan, modulation c, chitin and cellulose
synthases: a superfamily of cell-wall associated carbohydrate
polymerizing enzymes
Unpublished
REFERENCE 2 (bases 1 to 2109)
AUTHORS Martinez-Duncker,I., Oriol,R. and Mollicone,R.
JOURNAL Direct Submission
TITLE Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation
Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif
94807, France
JOURNAL
FEATURES
source
Location/Qualifiers
1..2109
/organism="Papio anubis"
/mol_type="mRNA"
/db_xref="taxon:9555"
1..2109
/gene="HAS1"
36..1787
/gene="HAS1"
/feature="cell wall associated protein; produces hyaluronan
acid chains"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAR25554.1"
/db_xref="GI:38607342"
/translation="WTQRTDTPKTPAARRCCSGLARRVLTIAFALLILGLMTWAYAAGV
PLASDRYGLLAFGLYGAFLSAHLAQSLPAYLEHRRVAAARSLAKGPDLAATARSV
ALTISAYQEDPAYLRQCLVSARALLYPHRLRLVMVDGNRPEDLYMVMDFREVFADDE
PATYVWDGNYPHPWEPAEATGAGVEGAYREVEADPGRLAVALVTRRCVCVQARWG
GKREVMYAFKALGSDVDYQVCDSDTRLDPMALLELVQLDEDPVGVAGGVDVRLIN
PLDSWVSFLSLRYWFAFNVACOSYFHCVSCI SGPLGLYRNLLQQFLKAWYNQKF
LGTHCTFGDDRHLTNRMLSMGYATKTSRSRCYSETPSSFLRWLSQOTRWSKSYFREW
LYNALWHRHHAWMTYEAVVSGLFFEFVAATVIRLFYAGRPWALLVLLVCVQVALAK
AAFAAWLRGCVRMVLLSYAPLYMCGLLPAKFLALVTMNSGNGTSGRKKLAANYVPV
LPALWALLLLGLLARSVAQEARADWSGFSRAAEAYHLLAAGAGAYYVYVWMLTIYVY
GVRRLCRRRSGGYRVQV"

ORIGIN
Query Match 87.0%; Score 17.4; DB 9; Length 2109;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
|||||
Db 1149 AGGGTCTCTCGAGTAGCA 1131

RESULT 19
AC108651_3

```



## WPCOMMENT

Sequence split into 4 fragments LOCUS AC108651 Accession AC108651  
Fragment Name Begin End

AC108651\_0 1 110000

AC108651\_1 100001 210000

AC108651\_2 200001 310000

AC108651\_3 300001 372955

Continuation (4 of 4) of AC108651 from base 300001 (AC108651 Rattus norvegicus clone CH2)

## Query Match

Best Local Similarity 87.0%; Score 17.4; DB 2; Length 72955;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCA 19

|||||

Db 20251 AGGGCGTCTCTGAATAGCA 20269

## RESULT 20

## AC137058

LOCUS AC137058 185623 bp DNA linear HTG 21-FEB-2003

DEFINITION Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered

## ACCESSION

AC137058

VERSION AC137058.3 GI:28460766

KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.

SOURCE Papio anubis (olive baboon)

## ORGANISM

Papio anubis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopitheciinae; Papio.

1 (bases 1 to 185623)

Akter.N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,

Benjamin.B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,

Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,

Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,

Laric,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,

Maguiliag,C., Masello,C., Maskeri,B., McDowell,J.,

Parquillan,C., Pearson,R., Portnoy,M.E., Prasad,A.,

Reddi-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,

Standipop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,

Wecherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 185623)

Green,E.D.

Direct Submission

Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717

Grovetmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 185623)

Green,E.D.

Direct Submission

Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717

Grovetmont Circle, Gaithersburg, MD 20877, USA

On Feb 21, 2003 this sequence version replaced gi:27476124.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc\_zoonhgri.nih.gov

----- Project Information

Center project name: dtj

Center clone name: 126M05

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out

gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 183092 bases at least Q40

Consensus quality: 183851 bases at least Q30

Consensus quality: 184434 bases at least Q20

Insert size: 160000; agarose-fp

Insert size: 184723; sum-of-contigs

Quality coverage: 14.13x in Q20 bases; agarose-fp

Quality coverage: 12.24x in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 10 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 9984: contig of 9984 bp in length

\* 9985 10084: gap of unknown length

\* 10085 49805: contig of 39721 bp in length

\* 49806 49905: gap of unknown length

\* 49906 66936: contig of 17031 bp in length

\* 66937 67036: gap of unknown length

\* 67037 83604: contig of 16568 bp in length

\* 83605 83704: gap of unknown length

\* 83705 108246: contig of 24542 bp in length

\* 108247 112021: contig of 3675 bp in length

\* 112022 112121: gap of unknown length

\* 112122 134817: contig of 22696 bp in length

\* 134818 134917: gap of unknown length

\* 134918 147290: contig of 12373 bp in length

\* 147291 147390: gap of unknown length

\* 147391 185337: contig of 37947 bp in length

\* 185338 185437: gap of unknown length

\* 185438 185623: contig of 186 bp in length.

## FEATURES

## source

1. 185623

/organism="Papio anubis"

/mol\_type="genomic DNA"

/db\_xref="taxon:9555"

/clone="RP41-126M5"

/clone\_lib="RP41"

1. 15263

/note="clone overlaps with GenBank Accession Number

AC130273 clone RP41-30706 (center project name deu)"

1. .9984

/note="assembly\_fragment

clone end:T7

vector\_side:left"

10085..49805

/note="assembly\_fragment"

49906..66936

/note="assembly\_fragment"

67037..83604

/note="assembly\_fragment"

83705..108246

/note="assembly\_fragment"

108347..112021

/note="assembly\_fragment"

112122..134817

/note="assembly\_fragment"

134918..147290

/note="assembly\_fragment"

147391..185337

/note="assembly\_fragment"

147391..185337

/note="assembly\_fragment"

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

```

misc_feature      185438..185623
                  /note="assembly_fragment
clone_end:SP6
vector_side:right"

ORIGIN
Query Match      87.0%; Score 17.4; DB 2; Length 185623;
Best Local Similarity 94.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCA 19
    |||||
Db 131762 AGGGTGTCTCTGAGTAGCA 131780

RESULT 21
AC108317
LOCUS
DEFINITION
AC108317
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Escoto-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowitz,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louleseg,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackelmech,O., Okwuonu G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Polindexter,A., Popovic,D., Primus,E., Fu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanai,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,P., Zhang,J., Zhou,X., Zhou,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 189862)  
Worley,K.C.  
Direct Submission  
Submitted (27-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 189862)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Sep 14, 2002 this sequence version replaced gi:21737603.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPXZ  
Center clone name: CH230-255E18  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 174388 bases at least Q40  
Consensus quality: 176023 bases at least Q30  
Consensus quality: 177229 bases at least Q20  
Estimated insert size: 177068; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 188739: contig of 188739 bp in length  
\* 188740 188839: gap of unknown length  
\* 188840 189862: contig of 1023 bp in length.  
Location/Qualifiers  
1..189862  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-255E18"  
1..1317  
/note="wgs\_contig"

FEATURES  
source  
misc\_feature  
ORIGIN  
Query Match 87.0%; Score 17.4; DB 2; Length 189862;  
Best Local Similarity 94.7%; Pred. No. 2.7e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
|||||  
Db 24219 GGGCTTCTGAGTAGCAG 24237

RESULT 22  
AL593846  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AL593846 213263 bp DNA linear ROD 18-SEP-2003  
Mouse DNA sequence from clone Rp23-171H16 on chromosome 11,  
complete sequence.  
AL593846  
AL593846.15 GI:20792584  
HTG.  
Mus musculus (house mouse)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 213263)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Tracey, A.

Direct Submission  
Submitted (17-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On May 14, 2002 this sequence version replaced gi:19848059.

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep -----

Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk

-----  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

RP23-171H16 is from the RPCI-23 Mouse BAC Library

constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a phred quality of at least 30.

## FEATURES

source

1. 213263  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone\_lib="RPCI-23"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 213263;  
Best Local Similarity 94.7%; Pred. No. 2.6e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
|||||  
Db 197418 AGGGCGTCTCTGAGTAGCA 197436

RESULT 23  
AC116203/c

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC116203 257325 bp DNA linear HTG 13-MAY-2003  
Rattus norvegicus clone CH230-148U11, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 4 unordered pieces.  
AC116203 GI:30579280  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 257325)  
Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biewalt, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorjis, E., Geer, K., Gill, R., Grady, M., Guerra, W.,  
Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensu, H., Lu, X., Ma, J.,  
Lorensu, H., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,  
Mawney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puzos, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajis, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,  
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,  
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 257325)

Worley, K. C.

Direct Submission

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE



\* 16310 17904: contig of 1595 bp in length  
 \* 17905 18004: gap of unknown length  
 \* 18005 19316: contig of 1312 bp in length  
 \* 19317 19416: gap of unknown length  
 \* 19417 21238: contig of 1822 bp in length  
 \* 21239 21338: gap of unknown length  
 \* 21339 25002: contig of 4184 bp in length  
 \* 25003 25602: gap of unknown length  
 \* 25603 27919: contig of 2317 bp in length  
 \* 27920 28019: gap of unknown length  
 \* 28020 31424: contig of 3405 bp in length  
 \* 31425 31524: gap of unknown length  
 \* 31525 34478: contig of 2954 bp in length  
 \* 34479 34578: gap of unknown length  
 \* 34579 39184: contig of 4606 bp in length  
 \* 39185 39284: gap of unknown length  
 \* 39285 43661: contig of 4377 bp in length  
 \* 43662 43761: gap of unknown length  
 \* 43762 49666: contig of 6205 bp in length  
 \* 49667 50667: gap of unknown length  
 \* 50668 56743: contig of 6677 bp in length  
 \* 56744 56843: gap of unknown length  
 \* 56844 64029: contig of 7186 bp in length  
 \* 64030 64129: gap of unknown length  
 \* 64130 71006: contig of 6877 bp in length  
 \* 71007 71106: gap of unknown length  
 \* 71107 79434: contig of 8328 bp in length  
 \* 79435 79534: gap of unknown length  
 \* 79535 92429: contig of 12895 bp in length  
 \* 92430 92529: gap of unknown length  
 \* 92530 111247: contig of 18718 bp in length  
 \* 111248 111347: gap of unknown length  
 \* 111348 122768: contig of 11421 bp in length  
 \* 122769 122868: gap of unknown length  
 \* 122869 141721: contig of 18853 bp in length  
 \* 141722 141821: gap of unknown length  
 \* 141822 156874: contig of 15053 bp in length  
 \* 156875 156974: gap of unknown length  
 \* 156975 176989: contig of 20015 bp in length  
 \* 176990 177089: gap of unknown length  
 \* 177090 197198: contig of 20109 bp in length  
 \* 197199 197298: gap of unknown length  
 \* 197299 215257: contig of 17959 bp in length  
 \* 215258 215357: gap of unknown length  
 \* 215358 230320: contig of 14963 bp in length  
 \* 230321 230420: gap of unknown length  
 \* 230421 256851: contig of 26431 bp in length  
 \* 256852 256951: gap of unknown length  
 \* 256952 293184: contig of 36233 bp in length.

# FEATURES

source  
 1..293184  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-182D19"  
 /clone\_lib="RPCI mouse BAC library 23"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 2; Length 293184;  
 Best Local Similarity 94.7%; Pred. No. 2.4e+02;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCA 19  
 |||||  
 Db 277530 AGGGCGTCTCTGAGTAGCA 277548

RESULT 25  
 AX898222  
 LOCUS AX898222 278 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 14085 from Patent EP1033401.  
 ACCESSION AX898222  
 VERSION AX898222.1 GI:40053135

# KEYWORDS

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1  
 AUTHORS Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.  
 TITLE Expressed sequence tags and encoded human proteins  
 JOURNAL Patent: EP 1033401-A 14085 06-SEP-2000;  
 Genset (FR)

## FEATURES

source  
 1..278  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 278;  
 Best Local Similarity 90.0%; Pred. No. 2.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||  
 Db 113 AGGGCATCTCTGAGTTGCAG 132

## RESULT 26

BD033755  
 LOCUS BD033755 278 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Sequence tag and encoded human protein.  
 ACCESSION BD033755  
 VERSION BD033755.1 GI:22575497  
 KEYWORDS JP 2001269182-A/10001.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 278)  
 AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.  
 TITLE Sequence tag and encoded human protein  
 JOURNAL Patent: JP 2001269182-A 10001 02-OCT-2001;  
 GENSET

## COMMENT

OS Homo sapiens (human)  
 PN JP 2001269182-A/10001  
 PD 02-OCT-2001  
 PR 24-FEB-2000 JP 2000118773  
 PR 26-FEB-1999 US 60/122487  
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, BIMERIC DUCLAIR, JEAN YVES  
 PI JORDAN  
 PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC  
 C12N5/10.  
 PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC  
 G06F15/40  
 CC

## FEATURES

source  
 1..278  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 278;  
 Best Local Similarity 90.0%; Pred. No. 2.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||  
 Db 113 AGGGCATCTCTGAGTTGCAG 132

## RESULT 27

AX886427

LOCUS AX886427 313 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 2290 from Patent EP1033401.  
ACCESSION AX886427  
VERSION AX886427.1 GI:40043577  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.  
AUTHORS Expressed sequence tags and encoded human proteins  
TITLE Patent: EP 1033401-A 2290 06-SEP-2000;  
JOURNAL Genset (FR)  
FEATURES  
source  
1..313  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
103..>312  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAP00565.1"  
/db\_xref="GI:40043578"  
/translation="MKPKGISLQHLNDYQRTQKXGHNSEPPRRRQCCPCYPCHT  
QRLPLCLLVLTITVIGRPLVCFLEP"  
ORIGIN  
Query Match 84.0%; Score 16.8; DB 6; Length 313;  
Best Local Similarity 90.0%; Pred. No. 2.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
Db 113 AGGGCATCTCTGAGTTGCAG 132  
RESULT 28  
LOCUS BD026037 313 bp DNA linear PAT 27-AUG-2002  
DEFINITION Sequence tag and encoded human protein.  
ACCESSION BD026037  
VERSION BD026037.1 GI:22567260  
KEYWORDS JP 2001269182-A/2283.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 313)  
REFERENCE Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.  
AUTHORS Sequence tag and encoded human protein  
TITLE Patent: JP 2001269182-A 2283 02-OCT-2001;  
JOURNAL GENSET  
COMMENT OS Homo sapiens (human)  
PN JP 2001269182-A/2283  
PD 02-OCT-2001  
PF 24-FEB-2000 JP 2000118773  
PR 26-FEB-1999 US 60/122487  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES  
PJ JORDAN  
PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC  
C12N5/10,  
PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC  
G06F15/40  
CC Key Location/Qualifiers  
FH CDS 103..312.  
FT CDS  
1..313  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
FEATURES  
source  
ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 313;  
Best Local Similarity 90.0%; Pred. No. 2.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
Db 113 AGGGCATCTCTGAGTTGCAG 132  
RESULT 29  
LOCUS AF367431 465 bp mRNA linear PRI 24-MAY-2001  
DEFINITION Homo sapiens clone 0484e082 placental protein 13-like mRNA, partial  
cds.  
ACCESSION AF367431  
VERSION AF367431.1 GI:14194180  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 465)  
REFERENCE Yang,Q.-S., Xie,Y. and Mao,Y.M.  
AUTHORS Cloning and expression pattern study of a novel human galectin  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 465)  
AUTHORS Yang,Q.-S., Xie,Y. and Mao,Y.M.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2001) Research Institute of Genetics, School of  
Life Science, 220 Handan Rd., Shanghai 200433, China  
FEATURES  
source  
1..465  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.1"  
/clone="0484e082"  
/tissue\_type="brain"  
/dev\_stage="18 week fetus"  
224..>465  
/codon\_start=1  
/product="placental protein 13-like"  
/protein\_id="AAK56285.1"  
/db\_xref="GI:14194181"  
/translation="MSSLVPVPTLPVSLPVGSCVITGPILTFVKDPLEVNFYTG  
DEDSDIATQFRLHFGHPAIMNSCVFGIWRYEKCY"  
CDS  
Query Match 84.0%; Score 16.8; DB 9; Length 465;  
Best Local Similarity 90.0%; Pred. No. 2.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
Db 127 AGGGCATCTCTGAGTTGCAG 146  
RESULT 30  
LOCUS AF367432 529 bp mRNA linear PRI 24-MAY-2001  
DEFINITION Homo sapiens clone 0484e083 placental protein 13-like mRNA, partial  
cds.  
ACCESSION AF367432  
VERSION AF367432.1 GI:14194182  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 529)  
REFERENCE Yang,Q.-S., Xie,Y. and Mao,Y.M.  
AUTHORS

TITLE Cloning and expression pattern study of a novel human galectin  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 529)  
AUTHORS Yang, Q.-S., Xie, Y. and Mao, Y.M.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2001) Research Institute of Genetics, School of Life Science, 220 Handan Rd., Shanghai 200433, China

FEATURES  
source  
Location/Qualifiers  
1..529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.1"  
/clone="0484e083"  
/tissue\_type="brain"  
/dev\_stage="18 week fetus"  
224..>529  
/codon\_start=1  
/product="placental protein 13-like"  
/protein\_id="AAK56286.1"  
/db\_xref="GI:14194183"  
/translation="MSSLPLVPVPTLPVSLPVGSCVITGTPILTFVKDPQLEVNFVTGM"  
DEDSDIAQFRLHFGHPAIMNSCVFIWRYBEKCYLLPFEDGKPFELCIYVRHKYKV  
"

ORIGIN  
Query Match 84.0%; Score 16.8; DB 9; Length 529;  
Best Local Similarity 90.0%; Pred. No. 2.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CDS  
1 AGGGCGTCTCTGAGTAGCAG 20  
|||||  
127 AGGGCATCTCTGAGTTGCAG 146

RESULT 31  
AF367430 536 bp mRNA linear PRI 24-MAY-2001  
LOCUS Homo sapiens clone 0484e081 placental protein 13-like mRNA, partial cds  
DEFINITION AF367430.1 GI:14194178  
ACCESSION AF367430  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 536)  
AUTHORS Yang, Q.-S., Xie, Y. and Mao, Y.M.  
TITLE Cloning and expression pattern study of a novel human galectin  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 536)  
AUTHORS Yang, Q.-S., Xie, Y. and Mao, Y.M.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2001) Research Institute of Genetics, School of Life Science, 220 Handan Rd., Shanghai 200433, China

FEATURES  
source  
Location/Qualifiers  
1..536  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.1"  
/clone="0484e081"  
/tissue\_type="brain"  
/dev\_stage="18 week fetus"  
224..>536  
/codon\_start=1  
/product="placental protein 13-like"  
/protein\_id="AAK56284.1"  
/db\_xref="GI:14194179"  
/translation="MSSLPLVPVPTLPVSLPVGSCVITGTPILTFVKDPQLEVNFVTGM"

CDS  
1 AGGGCGTCTCTGAGTAGCAG 20  
|||||  
127 AGGGCATCTCTGAGTTGCAG 146

TITLE Cloning and expression pattern study of a novel human galectin  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 529)  
AUTHORS Yang, Q.-S., Xie, Y. and Mao, Y.M.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2001) Research Institute of Genetics, School of Life Science, 220 Handan Rd., Shanghai 200433, China

FEATURES  
source  
Location/Qualifiers  
1..529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.1"  
/clone="0484e083"  
/tissue\_type="brain"  
/dev\_stage="18 week fetus"  
224..>529  
/codon\_start=1  
/product="placental protein 13-like"  
/protein\_id="AAK56286.1"  
/db\_xref="GI:14194183"  
/translation="MSSLPLVPVPTLPVSLPVGSCVITGTPILTFVKDPQLEVNFVTGM"  
DEDSDIAQFRLHFGHPAIMNSCVFIWRYBEKCYLLPFEDGKPFELCIYVRHKYKV  
"

ORIGIN  
Query Match 84.0%; Score 16.8; DB 9; Length 529;  
Best Local Similarity 90.0%; Pred. No. 2.3e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CDS  
1 AGGGCGTCTCTGAGTAGCAG 20  
|||||  
127 AGGGCATCTCTGAGTTGCAG 146

RESULT 32  
BV037702 555 bp DNA linear STS 31-MAY-2003  
LOCUS S212P60157FH3.TO CZECHII/Ei Mus musculus STS genomic, sequence tagged site.  
DEFINITION BV037702  
ACCESSION BV037702.1 GI:31121597  
VERSION  
KEYWORDS STS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 555)  
AUTHORS Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.  
TITLE The mosaic structure of variation in the laboratory mouse genome  
JOURNAL Nature 420 (6915), 574-578 (2002)  
MEDLINE 22354684  
PUBMED 12466852  
COMMENT  
Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172521477  
Email: kersli@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 555  
Protocol:  
WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvImJ, C3H/HeJ and BALB/cByJ. The WGS reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated  
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.

FEATURES  
source  
Location/Qualifiers  
1..555  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="CZECHII/Ei"  
/db\_xref="taxon:10090"  
/map="9 12-492 120522219-120522688"  
/clone\_lib="CZECHII/Ei"  
<1..>555

STS  
ORIGIN  
Query Match 84.0%; Score 16.8; DB 11; Length 555;  
Best Local Similarity 90.0%; Pred. No. 2.2e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

CDS  
1 AGGGCGTCTCTGAGTAGCAG 20  
|||||  
386 AGGGTGTCTCTGTGTAGCAG 405

RESULT 33

```

AP267852      774 bp  mRNA  linear  PRI 11-JUN-2002
LOCUS      Homo sapiens placental protein 13-like protein mRNA, complete cds.
DEFINITION
ACCESSION  AF267852
VERSION    AF267852.1 GI:8745530
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 774)
AUTHORS    Yang,Q.S., Ying,K., Yuan,H.L., Chen,J.Z., Meng,X.F., Wang,Z.,
            Xie,Y. and Mao,Y.M.
TITLE      Cloning and expression of a novel human galectin cDNA,
            predominantly expressed in placenta(1)
JOURNAL    Biochim. Biophys. Acta 1574 (3), 407-411 (2002)
MEDLINE    21992454
PUBMED     11997112
REFERENCE   2 (bases 1 to 774)
AUTHORS    Mao,Y.M., Xie,Y., Yuan,H.L., Yang,Q.S., Wu,H., Xia,P. and Ying,K.
TITLE      Direct Submission
JOURNAL    Submitted (15-MAY-2000) Institute of Genetics, Fudan University,
            220 Handan Rd., Shanghai 200433, China
FEATURES             Location/Qualifiers
     source          1..774
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosomes="19"
                     /map="between D19S425 and D19S418"
                     224..643
                     /note="PPL13"
                     /codon_start=1
                     /product="placental protein 13-like protein"
                     /protein_id="AAF78953.1"
                     /db_xref="GI:8745531"
                     /translation="MSSLVPVPTLPVSLPVGCVLTGTPILTFVKDPQLEVNFYTM
                     DSDSDIAFQRLHFHGPALMNSCVFGIWRYSKCYLPFEDKPKPELCIYVRHKYKV
                     MVNGORIYNFAHRFPSPVSKMLQVFRDLSLTRLVLSID"
     CDS             224..643
                     /note="PPL13"
                     /codon_start=1
                     /product="placental protein 13-like protein"
                     /protein_id="AAF78953.1"
                     /db_xref="GI:8745531"
                     /translation="MSSLVPVPTLPVSLPVGCVLTGTPILTFVKDPQLEVNFYTM
                     DSDSDIAFQRLHFHGPALMNSCVFGIWRYSKCYLPFEDKPKPELCIYVRHKYKV
                     MVNGORIYNFAHRFPSPVSKMLQVFRDLSLTRLVLSID"
     ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 774;
Best Local Similarity 90.0%; Pred. No. 2.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 127 AGGGCATCTCTGAGTTGCAG 146
    ||||| ||||| ||||| |||||

RESULT 34
BV063966      795 bp  DNA  linear  STS 31-MAY-2003
LOCUS      S212P6739FB3.TO CZECHII/Bi Mus musculus STS genomic, sequence
DEFINITION
tagged site.
ACCESSION  BV063966
VERSION    BV063966.1 GI:31179761
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 795)
AUTHORS    Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
            Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
TITLE      The mosaic structure of variation in the laboratory mouse genome
JOURNAL    Nature 420 (6915), 574-578 (2002)
MEDLINE    22354684
PUBMED     12466852
COMMENT
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersll@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 795
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGCv3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.
FEATURES             Location/Qualifiers
     source          1..795
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="CZECHII/Bi"
                     /db_xref="taxon:10090"
                     /map="5 22-744 123891032-123891755"
                     /clone_lib="CZECHII/Bi"
                     <1..>795
     ORIGIN
Query Match      84.0%; Score 16.8; DB 11; Length 795;
Best Local Similarity 90.0%; Pred. No. 2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 283 AGGGCTTCTCTGAGGAGCAG 302
    ||||| ||||| ||||| |||||

RESULT 35
AK056640/c    3111 bp  mRNA  linear  PRI 30-JAN-2004
LOCUS      Homo sapiens cDNA FLJ32078 fis, clone OCBF1000192.
DEFINITION
ACCESSION  AK056640
VERSION    AK056640.1 GI:16552098
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1
            Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hirakawa,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hikota,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,

```



Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isoqai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

**TITLE** Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahori, K., Masuho, Y., Nagai, K. and Isoqai, T.

**JOURNAL** NEDO human cDNA sequencing project

**PUBLISHED** Unpublished

**REFERENCE** 3 (bases 1 to 3111)

**AUTHORS** Isoqai, T., Otsuki, T. and Sugiyama, T.

**JOURNAL** Direct Submission

**TITLE** Submitted (24-OCT-2001) Takao Isoqai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

**JOURNAL** NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI); cDNA library Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

**COMMENT**

**FEATURES** Location/Qualifiers

**source** 1..3111

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="OCBBF1000192"

/tissue\_type="brain"

/clone\_lib="OCBBF1"

/dev\_stage="fetus"

/note="cloning vector: pME18SFL3"

**ORIGIN**

Query Match 84.0%; Score 16.8; DB 9; Length 3111;

Best Local Similarity 90.0%; Pred. No. 1.5e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGGGCTCTCGTAGTCAG 20

|||||

Db 380 AGGGAGTCTCTAGAGACGAG 361

|||||

RESULT 36

AC006133

**LOCUS** 43859 bp DNA linear PRI 05-DEC-1998

**DEFINITION** Homo sapiens chromosome 19, cosmid R30692, complete sequence.

**ACCESSION** AC006133

**VERSION** AC006133.1 GI:3970929

**KEYWORDS** HTG.

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

**AUTHORS** 1 (bases 1 to 43859)

Lamerdin, J.E., McCreedy, P.M., Skowronski, E., Vasanathan, V., Burkhardt-Schultz, K.J., Gordon, L., Dias, J., Ramirez, M., Stiilwagen, S., Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Barnes, J., Danganan, L., Eriar, A., Christensen, M., Georgescu, A., Avila, J., Liu, S., Attix, C., Andreise, T., Trankheim, M., Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Kronmiller, B., Arellano, A., Sanders, C., Ow, D.,

**TITLE** Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V. Sequence analysis of a 2.3 Mb region in 19q13.1 containing the RYR gene

**JOURNAL** Unpublished

**REFERENCE** 2 (bases 1 to 43859)

**AUTHORS** Lamerdin, J.E.

**JOURNAL** Direct Submission

**TITLE** Submitted (05-DEC-1998) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA

**FEATURES** Location/Qualifiers

**source** 1..43859

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="19"

/map="19q13.1 between D19S220 and AKT2"

/clone="R30692"

/cell\_line="5HL2-B"

/clone\_lib="LL19NC03 R chromosome 19-specific cosmid library"

/note="Cosmid library constructed at LLNL from flow-sorted chromosomes from hybrid 5HL2-B, which carries chromosome 19 as its only human chromosome."

1878..1938

/rpt\_family="(CAA)n"

/complement(1947..2101)

/note="predicted exon, program: grail2exons\_human\_1.3, frame: 1, quality: good, score: 73.000"

repeat\_region

misc\_feature

complement(2611..2907)

/rpt\_family="AluSq"

/complement(3392..3483)

/note="predicted exon, program: grail2exons\_human\_1.3, frame: 1, quality: excellent, score: 92.000"

3628..4845

/rpt\_family="L1PA14"

4846..5140

/rpt\_family="AluJb"

5141..5936

/rpt\_family="L1PA14"

6328..6465

/rpt\_family="AluJo/FRAM"

/complement(6508..6702)

/rpt\_family="AluY"

6703..7517

/rpt\_family="L1PA13"

/complement(8575..8639)

/rpt\_family="(CA)n"

9297..9594

/rpt\_family="AluJo"

/complement(9947..10157)

/rpt\_family="MIR"

/complement(11543..11844)

/rpt\_family="AluSx"

/complement(13142..13250)

/note="predicted exon, program: grail2exons\_human\_1.3, frame: 0, quality: excellent, score: 96.000"

complement(14710..14797)

/note="predicted exon, program: grail2exons\_human\_1.3, frame: 1, quality: excellent, score: 75.000"

complement(15955..16394)

/rpt\_family="L1PA11"

/complement(16389..16538)

/rpt\_family="L1MG"

/complement(16533..17661)

/rpt\_family="L1PB1"

/complement(17692..23827)

/rpt\_family="L1PA2"

23871..24384

/rpt\_family="L1PA7"

/complement(24392..24418)

/rpt\_family="(TA)n"

24716..24994

/rpt\_family="AluJo"



```

* 6089 6188: gap of 100 bp
* 7044: contig of 856 bp in length
* 7045 7144: gap of 100 bp
* 8032: contig of 888 bp in length
* 8033 8132: gap of 100 bp
* 8133 9027: contig of 895 bp in length
* 9028 9127: gap of 100 bp
* 9128 10040: contig of 913 bp in length
* 10041 10140: gap of 100 bp
* 10141 11108: contig of 968 bp in length
* 11109 11208: gap of 100 bp
* 11209 12041: contig of 833 bp in length
* 12042 12141: gap of 100 bp
* 12142 13050: contig of 909 bp in length
* 13051 13150: gap of 100 bp
* 13151 14095: contig of 945 bp in length
* 14096 14195: gap of 100 bp
* 14196 15050: contig of 855 bp in length
* 15051 15150: gap of 100 bp
* 15151 16097: contig of 947 bp in length
* 16098 16197: gap of 100 bp
* 16198 17167: contig of 970 bp in length
* 17168 17267: gap of 100 bp
* 17268 18207: contig of 940 bp in length
* 18208 18307: gap of 100 bp
* 18308 19191: contig of 884 bp in length
* 19192 19291: gap of 100 bp
* 19292 20292: contig of 1001 bp in length
* 20293 20392: gap of 100 bp
* 20393 21236: contig of 844 bp in length
* 21237 21336: gap of 100 bp
* 21337 22266: contig of 930 bp in length
* 22267 22366: gap of 100 bp
* 22367 23345: contig of 973 bp in length
* 23346 24365: contig of 920 bp in length
* 24366 24465: gap of 100 bp
* 24466 25416: contig of 951 bp in length
* 25417 26511: contig of 995 bp in length
* 26512 27513: contig of 902 bp in length
* 27514 28575: contig of 962 bp in length
* 28576 29572: contig of 897 bp in length
* 29573 29672: gap of 100 bp
* 29673 30706: contig of 934 bp in length
* 30707 31662: contig of 956 bp in length
* 31663 31762: gap of 100 bp
* 31763 32704: contig of 942 bp in length
* 32705 32804: gap of 100 bp
* 32805 33651: contig of 847 bp in length
* 33652 33751: gap of 100 bp
* 33752 34685: contig of 934 bp in length
* 34686 34785: gap of 100 bp
* 34786 35685: contig of 900 bp in length
* 35686 35785: gap of 100 bp
* 35786 36754: contig of 969 bp in length
* 36755 37764: gap of 100 bp
* 37765 37864: contig of 910 bp in length
* 37865 38804: contig of 940 bp in length
* 38805 38904: gap of 100 bp
* 38905 39865: contig of 961 bp in length
* 39866 39965: gap of 100 bp
* 39966 40921: contig of 956 bp in length
* 40922 41021: gap of 100 bp
* 41022 41934: contig of 913 bp in length
* 41935 42034: gap of 100 bp
* 42035 43049: contig of 1015 bp in length
* 43050 43149: gap of 100 bp

```

```

* 43150 44107: contig of 958 bp in length
* 44108 44207: gap of 100 bp
* 44208 45136: contig of 929 bp in length.
FEATURES
    source
        1..45136
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="RP23-127117"
            /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN
Query Match      84.0%; Score 16.8; DB 2; Length 45136;
Best Local Similarity 90.0%; Pred. No. 7.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
DB 23870 AGGGCGACTCGGAGTAGCAG 23851
RESULT 38
AC016271/c
LOCUS AC016271 Homo sapiens clone RP11-19F12, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC016271.2 GI:9123891
ACCESSION AC016271
VERSION HTG; HTGS_PHASE0.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 80065)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-19F12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 80065)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boquelavkiy,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teschke,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wymann,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6467063.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3728
Center clone name: 19_F_12
* NOTE: This record contains 82 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows

```

\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1 903: contig of 903 bp in length  
\* 904 1003: gap of 100 bp  
\* 1004 1895: contig of 892 bp in length  
\* 1896 1995: gap of 100 bp  
\* 1996 2890: contig of 895 bp in length  
\* 2891 2990: gap of 100 bp  
\* 2991 3816: contig of 826 bp in length  
\* 3817 3916: gap of 100 bp  
\* 3917 4776: contig of 860 bp in length  
\* 4777 4876: gap of 100 bp  
\* 4877 5746: contig of 870 bp in length  
\* 5747 5846: gap of 100 bp  
\* 5847 6741: contig of 895 bp in length  
\* 6742 6841: gap of 100 bp  
\* 6842 7705: contig of 864 bp in length  
\* 7706 7805: gap of 100 bp  
\* 7806 8688: contig of 883 bp in length  
\* 8689 9660: contig of 872 bp in length  
\* 9661 9760: gap of 100 bp  
\* 9761 10661: contig of 901 bp in length  
\* 10662 10761: gap of 100 bp  
\* 10762 11625: contig of 864 bp in length  
\* 11626 11725: gap of 100 bp  
\* 11726 12619: contig of 894 bp in length  
\* 12620 12719: gap of 100 bp  
\* 12720 13621: contig of 902 bp in length  
\* 13622 13721: gap of 100 bp  
\* 13722 14582: contig of 861 bp in length  
\* 14583 14682: gap of 100 bp  
\* 14683 15537: contig of 855 bp in length  
\* 15538 15637: gap of 100 bp  
\* 15638 16517: contig of 880 bp in length  
\* 16518 16617: gap of 100 bp  
\* 16618 17504: contig of 887 bp in length  
\* 17505 17604: gap of 100 bp  
\* 17605 18479: contig of 875 bp in length  
\* 18480 18579: gap of 100 bp  
\* 18580 19458: contig of 879 bp in length  
\* 19459 19558: gap of 100 bp  
\* 19559 20428: contig of 870 bp in length  
\* 20429 20528: gap of 100 bp  
\* 20529 21424: contig of 896 bp in length  
\* 21425 21524: gap of 100 bp  
\* 21525 22423: contig of 899 bp in length  
\* 22424 22523: gap of 100 bp  
\* 22524 23367: contig of 844 bp in length  
\* 23368 23467: gap of 100 bp  
\* 23468 24369: contig of 902 bp in length  
\* 24370 24469: gap of 100 bp  
\* 24470 25324: contig of 855 bp in length  
\* 25325 25424: gap of 100 bp  
\* 25425 26267: contig of 843 bp in length  
\* 26268 26367: gap of 100 bp  
\* 26368 27235: contig of 868 bp in length  
\* 27236 27335: gap of 100 bp  
\* 27336 28210: contig of 875 bp in length  
\* 28211 28310: gap of 100 bp  
\* 28311 29200: contig of 890 bp in length  
\* 29201 29300: gap of 100 bp  
\* 29301 30174: contig of 874 bp in length  
\* 30175 30274: gap of 100 bp  
\* 30275 31162: contig of 888 bp in length  
\* 31163 31262: gap of 100 bp  
\* 31263 32105: contig of 843 bp in length  
\* 32106 32205: gap of 100 bp  
\* 32206 33111: contig of 906 bp in length  
\* 33112 33211: gap of 100 bp

\* 33212 34119: contig of 908 bp in length  
\* 34120 34219: gap of 100 bp  
\* 34220 35103: contig of 884 bp in length  
\* 35104 35203: gap of 100 bp  
\* 35204 36064: contig of 861 bp in length  
\* 36065 36164: gap of 100 bp  
\* 36165 37033: contig of 869 bp in length  
\* 37034 37133: gap of 100 bp  
\* 37134 38021: contig of 888 bp in length  
\* 38022 38121: gap of 100 bp  
\* 38122 39118: contig of 897 bp in length  
\* 39119 39118: gap of 100 bp  
\* 39119 40034: contig of 916 bp in length  
\* 40035 40134: gap of 100 bp  
\* 40135 41047: contig of 913 bp in length  
\* 41048 41147: gap of 100 bp  
\* 41148 42068: contig of 921 bp in length  
\* 42069 42168: gap of 100 bp  
\* 42169 43048: contig of 880 bp in length  
\* 43049 43148: gap of 100 bp  
\* 43149 44027: contig of 873 bp in length  
\* 44028 44127: gap of 100 bp  
\* 44128 44989: contig of 862 bp in length  
\* 44990 45089: gap of 100 bp  
\* 45090 45956: contig of 867 bp in length  
\* 45957 46056: gap of 100 bp  
\* 46057 46940: contig of 884 bp in length  
\* 46941 47040: gap of 100 bp  
\* 47041 47939: contig of 899 bp in length  
\* 47940 48039: gap of 100 bp  
\* 48040 48930: contig of 791 bp in length  
\* 48931 48930: gap of 100 bp  
\* 48931 49793: contig of 863 bp in length  
\* 49794 49893: gap of 100 bp  
\* 49894 50799: contig of 906 bp in length  
\* 50800 50899: gap of 100 bp  
\* 50900 51768: contig of 869 bp in length  
\* 51769 51868: gap of 100 bp  
\* 51869 52723: contig of 855 bp in length  
\* 52724 52823: gap of 100 bp  
\* 52824 53686: contig of 863 bp in length  
\* 53687 53786: gap of 100 bp  
\* 53787 54671: contig of 885 bp in length  
\* 54672 54771: gap of 100 bp  
\* 54772 55556: contig of 885 bp in length  
\* 55557 55756: gap of 100 bp  
\* 55757 56644: contig of 888 bp in length  
\* 56645 56744: gap of 100 bp  
\* 56745 57637: contig of 893 bp in length  
\* 57638 57737: gap of 100 bp  
\* 57738 58616: contig of 879 bp in length  
\* 58617 58716: gap of 100 bp  
\* 58717 59555: contig of 839 bp in length  
\* 59556 59655: gap of 100 bp  
\* 59656 60555: contig of 900 bp in length  
\* 60556 61541: contig of 886 bp in length  
\* 61542 61641: gap of 100 bp  
\* 61642 62547: contig of 906 bp in length  
\* 62548 62647: gap of 100 bp  
\* 62648 63491: contig of 844 bp in length  
\* 63492 63591: gap of 100 bp  
\* 63592 64444: contig of 853 bp in length  
\* 64445 64544: gap of 100 bp  
\* 64545 65520: contig of 876 bp in length  
\* 65521 65520: gap of 100 bp  
\* 65521 66396: contig of 876 bp in length  
\* 66397 66497: gap of 100 bp  
\* 66497 67376: contig of 880 bp in length  
\* 67377 67476: gap of 100 bp  
\* 67477 68377: contig of 901 bp in length  
\* 68378 68477: gap of 100 bp  
\* 68478 69360: contig of 883 bp in length

```

* 69361 69460: gap of 100 bp
* 69461 70336: contig of 876 bp in length
* 70337 70436: gap of 100 bp
* 70437 71288: contig of 852 bp in length

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 2; Length 80065;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 77537 AGGGAGTCTCTGAGAGCAG 77518

RESULT 39
CR382131_11/c
WPCOMMENT
Sequence split into 43 fragments LOCUS CR382131 Accession CR382131
Fragment Name Begin End
CR382131_00 1 110000
CR382131_01 100001 210000
CR382131_02 200001 310000
CR382131_03 300001 410000
CR382131_04 400001 510000
CR382131_05 500001 610000
CR382131_06 600001 710000
CR382131_07 700001 810000
CR382131_08 800001 910000
CR382131_09 900001 1010000
CR382131_10 1000001 1110000
CR382131_11 1100001 1210000
CR382131_12 1200001 1310000
CR382131_13 1300001 1410000
CR382131_14 1400001 1510000
CR382131_15 1500001 1610000
CR382131_16 1600001 1710000
CR382131_17 1700001 1810000
CR382131_18 1800001 1910000
CR382131_19 1900001 2010000
CR382131_20 2000001 2110000
CR382131_21 2100001 2210000
CR382131_22 2200001 2310000
CR382131_23 2300001 2410000
CR382131_24 2400001 2510000
CR382131_25 2500001 2610000
CR382131_26 2600001 2710000
CR382131_27 2700001 2810000
CR382131_28 2800001 2910000
CR382131_29 2900001 3010000
CR382131_30 3000001 3110000
CR382131_31 3100001 3210000
CR382131_32 3200001 3310000
CR382131_33 3300001 3410000
CR382131_34 3400001 3510000
CR382131_35 3500001 3610000
CR382131_36 3600001 3710000
CR382131_37 3700001 3810000
CR382131_38 3800001 3910000
CR382131_39 3900001 4010000
CR382131_40 4000001 4110000
CR382131_41 4100001 4210000
CR382131_42 4200001 4224103
Continuation (12 of 43) of CR382131 from base 1100001 (CR382131 Yarrowia lipolytica chrC)

Query Match
Best Local Similarity 84.0%; Score 16.8; DB 8; Length 110000;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 99278 ACGGCGTCTCTGAGTAGCGG 99259

```

```

RESULT 40
AC106708
LOCUS
DEFINITION
Homo sapiens 3 BAC RP11-392A22 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
AC106708
AC106708.7 GI:21392433
HTG.
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 127109)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C.,
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouch, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lien, C., Liu, J., Liu, W.,
Louisleged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, Z., Mitchell, T., Mohabbat, K.,
Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,
Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,
Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N.,
Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,
Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M.,
Roias, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S.,
Scott, G., Shen, H., Shoohtari, N., Sison, I., Sodergren, E.,
Sonaiki, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,
Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,
Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, L.,
Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,
Warren, R., Washington, C., Watlington, S., Williams, G.,
Williamson, A., Wlarczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,
Wu, Y.F., Zhou, J., Zorrilla, S., Zorrilla, S., Zorrilla, S., Weinstein, G. and
Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 127109)
Worley, K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 127109)
Worley, K.C.
Direct Submission
Submitted (07-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 127109)
Worley, K.C.
Direct Submission
Submitted (12-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

```

## COMMENT

Baylor Plaza, Houston, TX 77030, USA  
On Jun 12, 2002 this sequence version replaced gi:21328473.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:  
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## FEATURES

source  
1. .127109  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-392A22"  
misc\_feature  
1. .2000  
/notes="overlaps bases 145860. .147859 of clone AC092993"  
/function="clone overlap"  
repeat\_region  
1. .469  
/rpt\_family="LTR8A"  
complement(520. .1695)  
repeat\_region  
/rpt\_family="Tigger2"  
1696. .1756  
repeat\_region  
/rpt\_family="MIR"  
complement(3752. .4287)  
repeat\_region  
/rpt\_family="MER67C"  
complement(5740. .5785)  
repeat\_region  
/rpt\_family="L2"  
6165. .6476  
repeat\_region  
/rpt\_family="AluJc"  
complement(7009. .7144)  
repeat\_region  
/rpt\_family="MER3"  
7145. .7449  
repeat\_region  
/rpt\_family="AluJo"  
8218. .9173  
repeat\_region  
/rpt\_family="pTR5"  
9174. .9806  
repeat\_region  
/rpt\_family="LTR12"  
complement(9816. .10114)  
repeat\_region  
/rpt\_family="AluSx"  
complement(10355. .10396)  
repeat\_region  
/rpt\_family="MIR"  
11331. .11520

repeat\_region  
/rpt\_family="AluSg/x"  
12355. .12458  
repeat\_region  
/rpt\_family="MLT1L"  
14405. .14425  
repeat\_region  
/rpt\_family="GC\_rich"  
14791. .14821  
repeat\_region  
/rpt\_family="GC\_rich"  
complement(15997. .16301)  
repeat\_region  
/rpt\_family="AluY"  
complement(16453. .16742)  
repeat\_region  
/rpt\_family="AluJo"  
16748. .16780  
repeat\_region  
/rpt\_family="AT\_rich"  
17344. .17430  
repeat\_region  
/rpt\_family="AT\_rich"  
complement(17616. .17685)  
repeat\_region  
/rpt\_family="L2"  
18061. .18192  
repeat\_region  
/rpt\_family="MER91A"  
20378. .20402  
repeat\_region  
/rpt\_family="AT\_rich"  
20569. .20827  
repeat\_region  
/rpt\_family="CTCG)n"  
20714. .20836  
STS  
/standard\_name="89729"  
complement(22797. .22844)  
repeat\_region  
/rpt\_family="MIR"  
complement(22837. .22970)  
repeat\_region  
/rpt\_family="L2"  
24520. .24611  
repeat\_region  
/rpt\_family="MIR"  
24766. .24812  
repeat\_region  
/rpt\_family="GGA)n"  
complement(25517. .25721)  
repeat\_region  
/rpt\_family="LIME"  
complement(26143. .26253)  
repeat\_region  
/rpt\_family="MER94"  
26505. .26539  
repeat\_region  
/rpt\_family="CA)n"  
27978. .28054  
repeat\_region  
/rpt\_family="MLT1L"  
28289. .28656  
repeat\_region  
/rpt\_family="MLT1A2"  
29260. .29427  
repeat\_region  
/rpt\_family="MER5A"  
complement(29760. .29892)  
repeat\_region  
/rpt\_family="MIR"  
complement(30077. .30368)  
repeat\_region  
/rpt\_family="AluSx"  
30445. .30783  
repeat\_region  
/rpt\_family="L2"  
30784. .31147  
repeat\_region  
/rpt\_family="THE1B"  
31148. .31596  
repeat\_region  
/rpt\_family="L2"  
31621. .32138

Query Match 84.0%; Score 16.8; DB 9; Length 127109;  
Best Local Similarity 90.0%; Pred. No. 5.8e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
|||||  
DB 53952 AGGGCTTCTCTGAGGAGCAG 53971

## RESULT 41

AL731701/c AL731701 128198 bp DNA linear ROD 20-AUG-2002  
DEFINITION Mouse DNA sequence from clone RP23-140L6 on chromosome X, complete  
sequence.  
ACCESSION AL731701  
VERSION AL731701.8 GI:22449772

**KEYWORDS**  
**SOURCE**  
**ORGANISM**  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 128198)  
**REFERENCE**  
**AUTHORS**  
 Heath, P.  
**TITLE**  
**JOURNAL**  
 Direct Submission  
 Submitted (20-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Aug 22, 2002 this sequence version replaced gi:21912692.  
**COMMENT**  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 -----  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; Sw:  
 SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-140L6 is  
 from the RPCI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACE3.6  
**FEATURES**  
 source  
 1..128198  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-140L6"  
 /clone\_lib="RPCI-23"  
**ORIGIN**  
 Query Match 84.0%; Score 16.8; DB 10; Length 128198;  
 Best Local Similarity 90.0%; Pred. No. 5.8e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 ||||| ||||| ||||| ||||| |||||  
 Db 10533 AGGGCGACTCGGAGTAGCAG 10514  
 RESULT 42  
 AC026671  
**LOCUS**  
**DEFINITION**  
 Homo sapiens 3 BAC RP11-91M9 (Roswell Park Cancer Institute Human  
 BAC Library) complete sequence.  
**ACCESSION**  
 AC026671  
**VERSION**  
 AC026671.18 GI:18921226  
**KEYWORDS**  
**SOURCE**  
 Homo sapiens (human)  
**ORGANISM**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 141674)  
**REFERENCE**  
**AUTHORS**  
 Muzny, D.M., Adams, C., Ali-Osman, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaraturunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carrroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, K.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, K., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,  
 Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,  
 Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
 Louised, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawney, E., McLeod, M.P., Meador, M.,  
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Rojass, A., Rojubokan, I., Roife, M., Ruiz, S., Savery, G., Scherer, S.,  
 Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E.,  
 Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,  
 Taber, P., Tanerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,  
 Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L.,  
 Vera, V., Villalobos, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,  
 Warren, R., Washington, C., Watlington, S., Williams, G.,  
 Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
 Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and  
 Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 141674)  
 Worley, K.C.  
 Direct Submission  
 Submitted (23-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 141674)  
 Worley, K.C.  
 Direct Submission  
 Submitted (26-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 141674)  
 Worley, K.C.  
 Direct Submission  
 Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Feb 26, 2002 this sequence version replaced gi:18644817.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)  
 CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.  
 ANNOTATION OF FEATURES:  
 STSs are identified using ePCR (Genome Res. 7:541-550) searches

of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

QUALSTAT-REPORT

Location/Qualifiers

```

1. .141674
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="3"
  /clone="RP11-91M9"
1. .2005

```

**מחיר**

repeat_region	/note="overlaps bases 139 /function="clone overlap" complement(236_616) /rpt_family="MER69"
repeat_region	complement(720_11140) /rpt_family="MER69"
repeat_region	complement(1287_1371) /rpt_family="MIR"
repeat_region	1401_1472 /rpt_family="MER5A"
repeat_region	/rpt_family="TATG)n" complement(2954_3069) /rpt_family="L2"
repeat_region	complement(3249_3420) /rpt_family="MER5A"
repeat_region	complement(3451_3538) /rpt_family="MER5B"
repeat_region	complement(3879_3957) /rpt_family="L2"
repeat_region	4314_4633 /rpt_family="AluSg1"
repeat_region	complement(6384_6597) /rpt_family="L2"
repeat_region	6598_6895 /rpt_family="AluSg1"
repeat_region	complement(6896_7058) /rpt_family="L2"
repeat_region	complement(7086_7397) /rpt_family="L2"
repeat_region	7996_8087 /rpt_family="MIR"
repeat_region	8381_8367 /rpt_family="MTR"
repeat_region	9088_9173 /rpt_family="L2"
repeat_region	complement(9712_10010) /rpt_family="AluSg"
repeat_region	10311_10333 /rpt_family="AluSg"
repeat_region	12072_12272 /note="AT rich"

repeat_region	13525 .13817 /rpt_family="L2"
repeat_region	13996 .14127 /rpt_family="L2"
repeat_region	complement(14715 .14974) /rpt_family="L1ME3"
repeat_region	14975 .15029 /rpt_family="CATATA"n"
repeat_region	16007 .16178 /rpt_family="LTR41"
repeat_region	16464 .16969 /rpt_family="LTR41"
repeat_region	complement(18618 .18796) /rpt_family="L1ME3A"
repeat_region	complement(18988 .19460) /rpt_family="L1M4"
repeat_region	19515 .19621 /rpt_family="MER91C"
repeat_region	complement(19634 .19782) /rpt_family="HAL1"
repeat_region	complement(19860 .20182) /rpt_family="L1M3C"
repeat_region	20236 .20265 /rpt_family="AT rich"
repeat_region	complement(20938 .21235) /rpt_family="AluY"
repeat_region	21862 .21886 /rpt_family="AT rich"
repeat_region	23464 .23516 /rpt_family="L1MA2"
repeat_region	23517 .23545 /rpt_family="AT rich"
repeat_region	23785 .24137 /rpt_family="THE1A"
repeat_region	25710 .26061 /rpt_family="THE1A"
repeat_region	26742 .26784 /rpt_family="TG"n"
repeat_region	complement(26819 .27044) /rpt_family="L2"
repeat_region	27553 .27608

```
Query Match      84.0%; Score 16.8; DB 9; Length 141674;
Best Local Similarity 90.0%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

**QY** 1 AGGGCGTCTCTGAGTAGCAG 20  
||| ||| ||| ||| |||  
**D6** 11754 AGGGATTCTCTGAGTAGCAG 11773

RESULT 43

AL627204	· 142447 bp	DNA linear	ROD 19-JUL-2002
LOCUS			
DEFINITION	Mouse DNA sequence from clone Rp23-118E21 on chromosome 4, complete sequence.		

ACCESSION	AL627204	GI:21953005	
VERSION	AL627204.7		
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 142447)		

Lead Direct	Submitted (19-JUL-2002)	Wellcome Trust
AUTHORS	Cambridgeshire, CB10 1SA, UK. E-mail en	
JOURNAL	humquerry@sanger.ac.uk Clone requests: c	
	On 24, 2002 this sequence version r	
COMMENT	----- Genome Center	
	Center: Wellcome Trust Sanger Institute	



Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-118E21 is from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBAC3.6.

## FEATURES

Location/Qualifiers  
1..142447  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-118E21"  
/clone\_lib="RPCI-23"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 142447;

Best Local Similarity 90.0%; Pred. No. 5.6e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20

Db 125628 AGGGCGTCTCTGAGGAGGAG 125647

## RESULT 44

AC142379/c

LOCUS AC142379 144791 bp DNA linear HTG 28-MAR-2003  
DEFINITION Rattus norvegicus clone CH230-428C22, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 52 unordered pieces.

ACCESSION AC142379

VERSION AC142379.1 GI:29336148

KEYWORDS HTG; HTGS\_PHASE1.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 144791)

Muzny,D,Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Caldeiron,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Kocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,

Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulsegged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangun,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaakemeleh,O., Okwuonu,G., Olanrunsgoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villagana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 144791)  
Worley,K.C.  
Submitted (28-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: KEJR  
Center clone name: CH230-428C22  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 109373 bases at least Q40  
Consensus quality: 118449 bases at least Q30  
Consensus quality: 124461 bases at least Q20  
Estimated insert size: 115919; sum-of-contigs estimation  
Quality coverage: 2x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1120: contig of 1120 bp in length  
\* 1121 1220: gap of unknown length  
\* 1221 2245: contig of 1025 bp in length

\* 2246 2345: gap of unknown length  
\* 2346 3765: contig of 1420 bp in length  
\* 3766 3865: gap of unknown length  
\* 3866 4902: contig of 1037 bp in length  
\* 4903 5002: gap of unknown length  
\* 5003 6331: contig of 1329 bp in length  
\* 6332 6431: gap of unknown length  
\* 6432 7644: contig of 1213 bp in length  
\* 7645 7744: gap of unknown length  
\* 7745 8790: contig of 1046 bp in length  
\* 8791 8890: gap of unknown length  
\* 8891 10769: contig of 1879 bp in length  
\* 10770 10859: gap of unknown length  
\* 10870 12171: contig of 1302 bp in length  
\* 12172 12271: gap of unknown length  
\* 12273 13631: contig of 1360 bp in length  
\* 13632 13731: gap of unknown length  
\* 13732 15082: contig of 1351 bp in length  
\* 15083 15182: gap of unknown length  
\* 15183 16360: contig of 1178 bp in length  
\* 16361 16460: gap of unknown length  
\* 16462 17642: contig of 1182 bp in length  
\* 17643 17742: gap of unknown length  
\* 17743 19814: contig of 2072 bp in length  
\* 19815 19914: gap of unknown length  
\* 19916 21109: contig of 1195 bp in length  
\* 21110 21209: gap of unknown length  
\* 21210 23322: contig of 2113 bp in length  
\* 23323 23422: gap of unknown length  
\* 23423 24487: contig of 1065 bp in length  
\* 24488 24587: gap of unknown length  
\* 24588 26099: contig of 1512 bp in length  
\* 26100 26199: gap of unknown length  
\* 26200 27327: contig of 1028 bp in length  
\* 27328 27328: gap of unknown length  
\* 27329 28359: contig of 1032 bp in length  
\* 28360 28459: gap of unknown length  
\* 28460 30141: contig of 1682 bp in length  
\* 30142 30241: gap of unknown length  
\* 30242 32324: contig of 2083 bp in length  
\* 32325 32424: gap of unknown length  
\* 32425 34042: contig of 1618 bp in length  
\* 34043 34142: gap of unknown length  
\* 34143 35445: contig of 1303 bp in length  
\* 35446 35545: gap of unknown length  
\* 35546 37447: contig of 1902 bp in length  
\* 37448 37548: gap of unknown length  
\* 37549 38795: contig of 1248 bp in length  
\* 38796 38895: gap of unknown length  
\* 38896 40985: contig of 2090 bp in length  
\* 40986 41085: gap of unknown length  
\* 41086 44415: contig of 3330 bp in length  
\* 44416 44515: gap of unknown length  
\* 44516 46750: contig of 2235 bp in length  
\* 46751 46850: gap of unknown length  
\* 46851 49596: contig of 2746 bp in length  
\* 49597 49696: gap of unknown length  
\* 49697 53101: contig of 3405 bp in length  
\* 53102 53201: gap of unknown length  
\* 53202 55283: contig of 2082 bp in length  
\* 55284 55383: gap of unknown length  
\* 55384 57687: contig of 2304 bp in length  
\* 57688 57787: gap of unknown length  
\* 57788 60536: contig of 2749 bp in length  
\* 60537 60636: gap of unknown length  
\* 60637 63422: contig of 2786 bp in length  
\* 63423 63522: gap of unknown length  
\* 63523 66737: contig of 3215 bp in length  
\* 66738 66837: gap of unknown length  
\* 66838 69284: contig of 2447 bp in length  
\* 69285 72049: contig of 2665 bp in length  
\* 72050 72149: gap of unknown length

\* 72150 74727: contig of 2578 bp in length  
\* 74728 74827: gap of unknown length  
\* 74828 78115: contig of 3288 bp in length  
\* 78116 78215: gap of unknown length  
\* 78216 83524: contig of 5309 bp in length  
\* 83525 83624: gap of unknown length  
\* 83625 88051: contig of 4427 bp in length  
\* 88052 88151: gap of unknown length  
\* 88152 89634: contig of 1483 bp in length  
\* 89635 89734: gap of unknown length  
\* 89735 95449: contig of 5715 bp in length  
\* 95450 95549: gap of unknown length  
\* 95550 99104: contig of 3555 bp in length  
\* 99105 99204: gap of unknown length  
\* 99205 104667: contig of 4863 bp in length  
\* 104668 104667: gap of unknown length  
\* 104669 109882: contig of 5715 bp in length  
\* 109883 109882: gap of unknown length  
\* 109884 114844: contig of 4862 bp in length  
\* 114845 114944: gap of unknown length  
\* 114945 120216: contig of 5272 bp in length  
\* 120217 120316: gap of unknown length  
\* 120317 127451: contig of 7135 bp in length  
\* 127452 127551: gap of unknown length  
\* 127552 135387: contig of 7836 bp in length  
\* 135388 135487: gap of unknown length  
\* 135488 144791: contig of 9304 bp in length.

## FEATURES Location/Qualifiers

Query Match 84.0%; Score 16.8; DB 2; Length 144791;  
Best Local Similarity 90.0%; Pred. No. 5.6e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGGCTCTCTGAGTAGCAG 20  
||||| ||||| ||||| |||||  
Db 94987 AGGGCTCTCTGAGTAGCAG 94968

## RESULT 45

AC073808 149727 bp DNA linear HTG 18-JUL-2000  
Mus musculus clone RP23-53L19, WORKING DRAFT SEQUENCE, 23 ordered  
pieces.  
AC073808 2 GI:9256809  
HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 149727)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Mouse  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 149727)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 18, 2000 this sequence version replaced gi:8810425.  
COMMENT -----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 1757507  
Center clone name: RPCI-23\_53L19  
-----  
Summary Statistics  
Consensus quality: 138531 bases at least Q40  
Consensus quality: 145586 bases at least Q30  
Consensus quality: 146846 bases at least Q20

Estimated insert size: 167000; agarose-fp estimation  
 Estimated insert size: 148677; sum-of-contigs estimation  
 Quality coverage: 6.17 in Q20 bases; agarose-fp estimation  
 Quality coverage: 6.93 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 23 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* by the accession number will be preserved.

\* 1 24435: contig of 24435 bp in length

\* 24436 24535: contig of 702 bp in length

\* 24536 25337: gap of unknown length

\* 25338 25337: gap of unknown length

\* 25338 27179: contig of 1842 bp in length

\* 27180 27279: gap of unknown length

\* 27280 37150: contig of 9871 bp in length

\* 37151 37250: gap of unknown length

\* 37251 43105: contig of 5855 bp in length

\* 43106 43205: gap of unknown length

\* 43206 44356: contig of 1051 bp in length

\* 44357 44356: gap of unknown length

\* 44357 57901: contig of 13545 bp in length

\* 57902 58001: gap of unknown length

\* 58002 71982: contig of 13981 bp in length

\* 71983 72082: gap of unknown length

\* 72083 79704: contig of 7622 bp in length

\* 79705 79804: gap of unknown length

\* 79805 84158: contig of 4254 bp in length

\* 84159 86603: gap of unknown length

\* 86604 86703: gap of unknown length

\* 86704 92822: contig of 6119 bp in length

\* 92823 105052: gap of unknown length

\* 105053 105152: gap of unknown length

\* 105153 106141: contig of 989 bp in length

\* 106142 106241: gap of unknown length

\* 106242 110743: contig of 4502 bp in length

\* 110744 110843: gap of unknown length

\* 110844 112176: contig of 1333 bp in length

\* 112177 112276: gap of unknown length

\* 112277 128768: contig of 16492 bp in length

\* 128769 128668: gap of unknown length

\* 128669 131960: contig of 3092 bp in length

\* 131961 132060: gap of unknown length

\* 132061 137483: contig of 5423 bp in length

\* 137484 137583: gap of unknown length

\* 137584 142994: contig of 5411 bp in length

\* 142995 143094: gap of unknown length

\* 143095 145122: contig of 2028 bp in length

\* 145123 145222: gap of unknown length

\* 145223 148984: contig of 3762 bp in length

\* 148985 149084: gap of unknown length

\* 149085 149727: contig of 643 bp in length.

#### FEATURES

source

1..149727  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-53119"  
 /clone\_lib="RPC1 mouse BAC library 23"

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 149727;  
 Best Local Similarity 90.0%; Pred. No. 5.5e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20

||||| ||||||| |||||

Db 68753 AGGGAGTCTCTGAGAGCAG 68772

#### RESULT 46

AC104782/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC104782 150010 bp DNA linear PRI 16-APR-2002  
 Homo sapiens BAC clone RP11-30G7 from 2, complete sequence.

AC104782

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)

----- Summary Statistics

Center project name: H\_NH0030G07

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis

MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenos, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-549H5, 2000 bp overlap; the clone sequenced to the right is RP11-292K15, 2000 bp overlap.

Actual start of this clone is at base position 161901 of RP11-549H5; actual end is at base position 18470 of RP11-292K15.

There are unresolved homopolymeric runs from 34720 to 34746 and 19405 to 19423.

#### FEATURES

##### source

##### Location/Qualifiers

```
1. .150010
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-30G7"
   /clone_lib="RPCI-11"
```

```
repeat_region 285..476
               /rpt_family="L2"
repeat_region 1269..1382
               /rpt_family="L1"
repeat_region 1396..1561
               /rpt_family="MER1_type"
repeat_region 1690..1911
               /rpt_family="L2"
repeat_region 2057..2169
               /rpt_family="MIR"
repeat_region 2382..2515
               /rpt_family="Alu"
repeat_region 2539..2854
               /rpt_family="Alu"
repeat_region 2855..3148
               /rpt_family="Alu"
repeat_region 3570..3729
               /rpt_family="L2"
repeat_region 3722..3889
               /rpt_family="MIR"
repeat_region 4843..5031
               /rpt_family="MIR"
repeat_region 5032..5329
               /rpt_family="Alu"
repeat_region 5330..5376
               /rpt_family="MIR"
repeat_region 5379..5411
               /rpt_family="CT-rich"
repeat_region 5412..5436
               /rpt_family="CA)n"
repeat_region 5737..5958
               /rpt_family="Alu"
repeat_region 6826..7108
               /rpt_family="Alu"
repeat_region 7115..7206
               /rpt_family="(GAAA)n"
repeat_region 7266..7394
               /rpt_family="Alu"
repeat_region 8707..9660
               /rpt_family="L1"
repeat_region 15756..16206
               /rpt_family="ERV1"
repeat_region 16207..16509
```

```
repeat_region 16510..16546
               /rpt_family="Alu"
repeat_region 16993..17300
               /rpt_family="Alu"
repeat_region 17936..18404
               /rpt_family="ERV1"
repeat_region 18567..18724
               /rpt_family="Alu"
repeat_region 18754..18773
               /rpt_family="(CAA)n"
repeat_region 19159..19255
               /rpt_family="L1"
repeat_region 19402..19424
               /rpt_family="(A)n"
repeat_region 20101..20407
               /rpt_family="Alu"
repeat_region 20459..21137
               /rpt_family="L1"
repeat_region 21269..21502
               /rpt_family="L1"
repeat_region 21624..21927
               /rpt_family="Alu"
repeat_region 21949..21973
               /rpt_family="(CAAAA)n"
repeat_region 21975..22200
               /rpt_family="L1"
repeat_region 22201..22518
               /rpt_family="Alu"
repeat_region 22519..23041
               /rpt_family="L1"
repeat_region 23047..23132
               /rpt_family="L1"
repeat_region 23179..23634
               /rpt_family="L1"
repeat_region 23637..23726
               /rpt_family="MER2_type"
repeat_region 23729..24237
               /rpt_family="MER2_type"
repeat_region 24243..25648
               /rpt_family="L1"
repeat_region 25808..25881
               /rpt_family="L1"
repeat_region 25882..25908
               /rpt_family="(TG)n"
repeat_region 25909..26454
               /rpt_family="L1"
repeat_region 26455..27618
               /rpt_family="L1"
repeat_region 27641..28259
               /rpt_family="L1"
repeat_region 28260..28561
               /rpt_family="Alu"
repeat_region 28562..29112
               /rpt_family="L1"
repeat_region 29113..29395
               /rpt_family="Alu"
repeat_region 29396..30357
               /rpt_family="L1"
```

Query Match 84.0%; Score 16.8; DB 9; Length 150010;  
Best Local Similarity 90.0%; Pred.No. 5.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGCTCTCTGAGTAGCAG 20  
||| ||||| |||||  
DB 65730 AGGGAGTCTCTGAGAAGCAG 65711

#### RESULT 47

AC133622  
LOCUS AC133622 150354 bp DNA linear HTG 16-SEP-2002  
DEFINITION Rattus norvegicus clone CH230-380C8, \*\*\* SEQUENCING IN PROGRESS

```

***, 61 unordered pieces.
AC133622
AC133622.1 GI:22901964
HTG: HTGS PHASE1.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 150354)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blythe, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgiadis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensheva, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinec, E.,
Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwaekelemeh, O., Okwuonu, G.,
Olanrunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,
Popovic, D., Primus, E., Pu, L., Puzos, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E.,
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Umami, K., Valas, R., Vera, V., Villaseana, D.,
Walton, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 150354)
Rat Genome Sequencing Consortium.
Rat Genome Sequencing
Submitted (16-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
-----
Center project name: KART

```

Center clone name: CH230-380C8  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 105186 bases at least Q40  
Consensus quality: 113495 bases at least Q30  
Consensus quality: 118387 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 61 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
as soon as it is available and the accession number will  
be preserved.

1 1052: contig of 1052 bp in length  
1152: gap of unknown length  
1153 2648: contig of 1496 bp in length  
1153 2649: gap of unknown length  
2649 2748: contig of 1325 bp in length  
2749 4074: contig of 1350 bp in length  
4074 4174: gap of unknown length  
4174 5523: contig of 1350 bp in length  
5524 5624: gap of unknown length  
5624 6938: contig of 1315 bp in length  
6939 7039: gap of unknown length  
7039 9211: contig of 2172 bp in length  
9211 9310: gap of unknown length  
9311 10861: contig of 1551 bp in length  
10862 10961: gap of unknown length  
10962 12573: contig of 1612 bp in length  
12574 12674: gap of unknown length  
12674 13810: contig of 1137 bp in length  
13811 13910: gap of unknown length  
13911 15321: contig of 1411 bp in length  
15322 15421: gap of unknown length  
15422 16800: contig of 1379 bp in length  
16801 16900: gap of unknown length  
16901 18199: contig of 1299 bp in length  
18200 18299: gap of unknown length  
18300 19568: contig of 1269 bp in length  
19569 19668: gap of unknown length  
19669 20808: contig of 1140 bp in length  
20809 20908: gap of unknown length  
20909 21977: contig of 1069 bp in length  
21978 22077: gap of unknown length  
22078 24333: contig of 2256 bp in length  
24334 24433: gap of unknown length  
24434 26234: contig of 1801 bp in length  
26235 26334: gap of unknown length  
26335 28821: contig of 2487 bp in length  
28822 30663: contig of 1742 bp in length  
30664 30763: gap of unknown length  
30764 32285: contig of 1522 bp in length  
32286 32385: gap of unknown length  
32386 34080: contig of 1695 bp in length  
34081 34180: gap of unknown length  
34181 35679: contig of 1499 bp in length  
35680 35779: gap of unknown length  
35780 37944: contig of 2165 bp in length  
37945 38044: gap of unknown length  
38045 40357: contig of 2313 bp in length  
40358 40457: gap of unknown length  
40458 43365: contig of 2908 bp in length  
43366 43465: gap of unknown length  
43466 45062: contig of 1597 bp in length  
45063 45162: gap of unknown length  
45163 46969: contig of 1807 bp in length  
46970 47069: gap of unknown length

```

* 47070 49637: contig of 2568 bp in length
* 49638 49737: gap of unknown length
* 49738 53643: contig of 3906 bp in length
* 53644 53743: gap of unknown length
* 53744 55353: contig of 1610 bp in length
* 55354 55453: gap of unknown length
* 55454 58176: contig of 2723 bp in length
* 58177 58276: gap of unknown length
* 58277 60684: contig of 2388 bp in length
* 60685 60784: gap of unknown length
* 60785 62799: contig of 2035 bp in length
* 62800 62899: gap of unknown length
* 62900 65808: contig of 2909 bp in length
* 65809 65908: gap of unknown length
* 65909 68722: contig of 2814 bp in length
* 68723 68822: gap of unknown length
* 68823 70990: contig of 2168 bp in length
* 70991 71090: gap of unknown length
* 71091 72739: contig of 1649 bp in length
* 72740 72839: gap of unknown length
* 72840 75263: contig of 2424 bp in length
* 75264 75363: gap of unknown length
* 75364 76380: contig of 1017 bp in length
* 76381 76480: gap of unknown length
* 76481 78500: contig of 2020 bp in length
* 78501 78600: gap of unknown length
* 78601 80836: contig of 2236 bp in length
* 80837 80936: gap of unknown length
* 80937 84694: contig of 3758 bp in length
* 84695 84794: gap of unknown length
* 84795 87092: contig of 2298 bp in length
* 87093 87192: gap of unknown length
* 87193 89597: contig of 2405 bp in length
* 89598 89697: gap of unknown length
* 89698 91676: contig of 1978 bp in length
* 91677 91775: gap of unknown length
* 91776 94065: contig of 2290 bp in length
* 94066 94165: gap of unknown length
* 94166 97567: contig of 3402 bp in length
* 97568 97667: gap of unknown length
* 97668 99869: contig of 2202 bp in length
* 99869 99969: gap of unknown length
* 99970 102460: contig of 2491 bp in length
* 102461 102560: gap of unknown length
* 102561 104482: contig of 1922 bp in length
* 104483 104582: gap of unknown length
* 104583 107242: contig of 2660 bp in length
* 107243 107342: gap of unknown length
* 107343 110239: contig of 2897 bp in length
* 110240 110339: gap of unknown length
* 110340 113361: contig of 3022 bp in length
* 113362 113461: gap of unknown length

```

Query Match 84.0%; Score 16.8; DB 2; Length 150354;  
 Best Local Similarity 90.0%; Pred. No. 5 se+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 66864 AGGGAGTCACTGAGTAGCAG 66883

```

```

RESULT 48
AC019258
LOCUS AC019258 155211 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-2M23, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC019258
VERSION AC019258.2 GI:7209488
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

# REFERENCE

1 (bases 1 to 155211)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens, clone RP11-2M23  
 Unpublished  
 2 (bases 1 to 155211)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,  
 Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
 Dearellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,  
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,  
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lander,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
 MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,  
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,  
 Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
 Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 8, 2000 this sequence version replaced gi:6649456.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: W18R  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L2672  
 Center clone name: 2\_M\_23  
 -----  
 \* NOTE: This record contains 177 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.  
 \* 1 756: contig of 756 bp in length  
 \* 757 856: gap of 100 bp  
 \* 857 1619: contig of 763 bp in length  
 \* 1620 1719: gap of 100 bp  
 \* 1720 2498: contig of 779 bp in length  
 \* 2499 2598: gap of 100 bp  
 \* 2599 3355: contig of 757 bp in length  
 \* 3356 3455: gap of 100 bp  
 \* 3456 4201: contig of 746 bp in length  
 \* 4202 4301: gap of 100 bp  
 \* 4302 5053: contig of 752 bp in length  
 \* 5054 5153: gap of 100 bp  
 \* 5154 5910: contig of 757 bp in length  
 \* 5911 6010: gap of 100 bp  
 \* 6011 6779: contig of 769 bp in length  
 \* 6780 6879: gap of 100 bp  
 \* 6880 7629: contig of 750 bp in length  
 \* 7630 7729: gap of 100 bp  
 \* 7730 8485: contig of 756 bp in length  
 \* 8486 8585: gap of 100 bp  
 \* 8586 9348: contig of 763 bp in length  
 \* 9349 9449: gap of 100 bp  
 \* 9449 10225: contig of 777 bp in length  
 \* 10226 10325: gap of 100 bp  
 \* 10326

## TITLE

## JOURNAL

## COMMENT

\* 10326 11083: contig of 758 bp in length  
\* 11084 11183: gap of 100 bp  
\* 11184 11964: contig of 781 bp in length  
\* 11965 12064: gap of 100 bp  
\* 12065 12830: contig of 766 bp in length  
\* 12831 12930: gap of 100 bp  
\* 12931 13692: contig of 762 bp in length  
\* 13693 13792: gap of 100 bp  
\* 13793 14557: contig of 765 bp in length  
\* 14558 14657: gap of 100 bp  
\* 14658 15412: contig of 755 bp in length  
\* 15413 15512: gap of 100 bp  
\* 15513 16282: contig of 770 bp in length  
\* 16283 16382: gap of 100 bp  
\* 16383 17131: contig of 749 bp in length  
\* 17132 17231: gap of 100 bp  
\* 17232 17980: contig of 749 bp in length  
\* 17981 18080: gap of 100 bp  
\* 18081 18939: contig of 759 bp in length  
\* 18940 19701: contig of 762 bp in length  
\* 19702 19801: gap of 100 bp  
\* 19802 20565: contig of 764 bp in length  
\* 20566 21415: gap of 100 bp  
\* 21416 22286: contig of 771 bp in length  
\* 22287 22386: gap of 100 bp  
\* 22387 23161: contig of 775 bp in length  
\* 23162 23262: contig of 761 bp in length  
\* 23263 24122: gap of 100 bp  
\* 24123 24872: contig of 750 bp in length  
\* 24873 25554: contig of 682 bp in length  
\* 25555 25754: gap of 100 bp  
\* 25755 26510: contig of 756 bp in length  
\* 26511 27369: contig of 759 bp in length  
\* 27370 27469: gap of 100 bp  
\* 27470 28209: contig of 740 bp in length  
\* 28210 28309: gap of 100 bp  
\* 28310 29077: contig of 768 bp in length  
\* 29078 29177: gap of 100 bp  
\* 29178 29924: contig of 747 bp in length  
\* 29925 30024: gap of 100 bp  
\* 30025 30796: contig of 772 bp in length  
\* 30797 30896: gap of 100 bp  
\* 30897 31655: contig of 759 bp in length  
\* 31656 31755: gap of 100 bp  
\* 31756 32502: contig of 747 bp in length  
\* 32503 32602: gap of 100 bp  
\* 32603 33369: contig of 767 bp in length  
\* 33370 33469: gap of 100 bp  
\* 33470 34233: contig of 764 bp in length  
\* 34234 34333: gap of 100 bp  
\* 34334 35181: gap of 100 bp  
\* 35082 35181: contig of 748 bp in length  
\* 35182 35951: contig of 770 bp in length  
\* 35952 36051: gap of 100 bp  
\* 36052 36920: contig of 769 bp in length  
\* 36921 36920: gap of 100 bp  
\* 36921 37694: contig of 774 bp in length  
\* 37695 37794: gap of 100 bp  
\* 37795 38563: contig of 769 bp in length  
\* 38564 38663: gap of 100 bp  
\* 38664 39427: contig of 764 bp in length  
\* 39428 39527: gap of 100 bp  
\* 39528 40285: contig of 758 bp in length  
\* 40286 40385: gap of 100 bp  
\* 40386 41148: contig of 763 bp in length  
\* 41149 41248: gap of 100 bp  
\* 41249 42007: contig of 759 bp in length

\* 42008 42107: gap of 100 bp  
\* 42108 42867: contig of 760 bp in length  
\* 42868 42967: gap of 100 bp  
\* 42968 43716: contig of 749 bp in length  
\* 43717 43816: gap of 100 bp  
\* 43817 44549: contig of 733 bp in length  
\* 44550 44649: gap of 100 bp  
\* 44650 45404: contig of 755 bp in length  
\* 45405 45504: gap of 100 bp  
\* 45505 46370: contig of 766 bp in length  
\* 46371 46370: gap of 100 bp  
\* 46371 47132: contig of 762 bp in length  
\* 47133 47232: gap of 100 bp  
\* 47233 48006: contig of 774 bp in length  
\* 48007 48106: gap of 100 bp  
\* 48107 48875: contig of 769 bp in length  
\* 48876 49735: contig of 760 bp in length  
\* 49736 49835: gap of 100 bp  
\* 49836 50585: contig of 750 bp in length  
\* 50586 51449: contig of 764 bp in length  
\* 51450 51549: gap of 100 bp  
\* 51550 52306: contig of 757 bp in length  
\* 52307 52406: gap of 100 bp  
\* 52407 53158: contig of 752 bp in length  
\* 53159 53258: gap of 100 bp  
\* 53259 54002: contig of 744 bp in length  
\* 54003 54102: gap of 100 bp  
\* 54103 54883: contig of 781 bp in length  
\* 54884 54983: gap of 100 bp  
\* 54984 55748: contig of 765 bp in length  
\* 55749 55848: gap of 100 bp  
\* 55849 56621: contig of 773 bp in length  
\* 56622 56721: gap of 100 bp  
\* 56722 57477: contig of 756 bp in length  
\* 57478 57577: gap of 100 bp  
\* 57578 58335: contig of 758 bp in length  
\* 58336 58435: gap of 100 bp  
\* 58436 59188: contig of 753 bp in length  
\* 59189 59288: gap of 100 bp  
\* 59289 60050: contig of 762 bp in length  
\* 60051 60150: gap of 100 bp  
\* 60151 60905: contig of 755 bp in length  
\* 60906 61005: gap of 100 bp

Query Match 84.0%; Score 16.8; DB 2; Length 155211;  
Best Local Similarity 90.0%; Pred. No. 5.5e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20  
Db 143745 ATGGCGTTCTGAGTAGCAG 143764

RESULT 49

AC128702/c

LOCUS

DEFINITION

AC128702

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AC128702 159431 bp DNA linear ROD 13-NOV-2003  
Mus musculus BAC clone RP24-239M3 from chromosome 9, complete  
sequence.  
AC128702  
HTG. AC128702.4 GI:28604184  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 159431)  
Dignan, G. and Cotton, M.  
The sequence of Mus musculus BAC clone RP24-239M3  
Unpublished (2001)  
2 (bases 1 to 159431)  
Wilson, R.

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 159431)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (22-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 159431)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (12-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 159431)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (28-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 159431)
Wilson,R.
Direct Submission
Submitted (13-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 28, 2003 this sequence version replaced gi:28316673.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
Center project name: M_BB0239M03
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1. .159431
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="9"
/map="9"
/clone="RP24-239M3"
/clone_lib="RPCI-24"
repeat_region 149. .265
/rpt_family="Alu"
repeat_region 737. .837
/rpt_family="B4"
repeat_region 901. .1235

FEATURES
source
repeat_region 1334. .1436
/rpt_family="MaLR"
repeat_region 1437. .1575
/rpt_family="B4"
repeat_region 1576. .1592
/rpt_family="B4"
repeat_region 1975. .2066
/rpt_family="B4"
repeat_region 2349. .2609
/rpt_family="ID"
repeat_region 2780. .3200
/rpt_family="ERVK"
repeat_region 3224. .3374
/rpt_family="B2"
repeat_region 3396. .3485
/rpt_family="B4"
repeat_region 3486. .3508
/rpt_family="Alu"
repeat_region 3826. .4037
/rpt_family="B2"
repeat_region 4285. .4468
/rpt_family="B2"
complement(4390. .4463)
/product="cRNA-Seq"
/note="Likely pseudogene (HMM Sc=37.66 / Sec struct
Sc=-5.63)"
repeat_region 4752. .4883
/rpt_family="Alu"
repeat_region 5535. .5677
/rpt_family="MIR"
repeat_region 6063. .6130
/rpt_family="B4"
repeat_region 9714. .9921
/rpt_family="B4"
repeat_region 10963. .11025
/rpt_family="B4"
repeat_region 11036. .11226
/rpt_family="B2"
trna 11041. .11113
/product="cRNA-Seq"
/note="Likely pseudogene (HMM Sc=31.58 / Sec struct
Sc=-4.91)"
repeat_region 11345. .11546
/rpt_family="B2"
repeat_region 11550. .11675
/rpt_family="Alu"
repeat_region 11750. .11849
/rpt_family="Alu"
repeat_region 11979. .12128
/rpt_family="Alu"
repeat_region 12213. .12337
/rpt_family="Alu"
repeat_region 13250. .13356
/rpt_family="Alu"
repeat_region 13589. .13753
/rpt_family="Alu"
repeat_region 14517. .14701
/rpt_family="B2"
repeat_region 15057. .15231
/rpt_family="B2"
repeat_region 15353. .15549
/rpt_family="B4"
repeat_region 15551. .15625
/rpt_family="MaLR"
repeat_region 15739. .15896
/rpt_family="Alu"
repeat_region 15991. .16100
/rpt_family="B4"
repeat_region 16183. .16290
/rpt_family="MaLR"
repeat_region 16415. .16620
/rpt_family="MaLR"

```



```

repeat_region 17554..17717
                /rpt_family="B2"
repeat_region 17718..17861
                /rpt_family="Alu"
repeat_region 17862..17886
                /rpt_family="B2"
repeat_region 17914..18037
                /rpt_family="Alu"
repeat_region 18048..18179
                /rpt_family="Alu"
repeat_region 18228..18368
                /rpt_family="B4"
repeat_region 18703..18838
                /rpt_family="B4"
repeat_region 19339..19479
                /rpt_family="Alu"
repeat_region 19511..19719
                /rpt_family="B2"
repeat_region 19875..20087
                /rpt_family="B2"
repeat_region 21583..21669
                /rpt_family="B4"
repeat_region 21685..21892
                /rpt_family="B2"
repeat_region 21914..22008
                /rpt_family="B2"
repeat_region 22065..22255
                /rpt_family="B4"
repeat_region 23167..23361
                /rpt_family="B2"
repeat_region 23516..23591
                /rpt_family="ID"
repeat_region 23596..23647
                /rpt_family="Alu"
repeat_region 23605..23663
                /rpt_family="B4"
repeat_region 23783..23870
                /rpt_family="Alu"

```

Query Match 84.0%; Score 16.8; DB 10; Length 159431;

Best Local Similarity 90.0%; Pred.No. 5.5e+02; Mismatches 0; Gaps 0;

Matches 18; Conservative 0; Indels 0; Gaps 0;

```

QY 1 AGGGCGCTCTCTGAGTAGCAG 20
Db 108672 AGGGTGTCTCTGTGAGTAGCAG 108653

RESULT 50
AC053534/c
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-745B12, WORKING DRAFT
AC053534
VERSION AC053534.3 GI:13270837
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 160872)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160872)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Mar 10, 2001 this sequence version replaced gi:11496398.
----- Genome Center -----

```

```

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: H_NH0745B12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 138905 bases at least Q40
Consensus quality: 146101 bases at least Q30
Consensus quality: 149339 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 156372; sum-of-contigs
Quality coverage: 2.92 in Q20 bases; agarose-fp
Quality coverage: 3.02 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1335: contig of 1335 bp in length
* 1336 1435: gap of unknown length
* 1436 2709: contig of 1274 bp in length
* 2710 2809: gap of unknown length
* 2810 4388: contig of 1579 bp in length
* 4389 4488: gap of unknown length
* 4489 5988: contig of 1500 bp in length
* 5989 6089: gap of unknown length
* 6090 8211: contig of 2022 bp in length
* 8212 9403: contig of 1193 bp in length
* 9404 9503: gap of unknown length
* 9504 10749: contig of 1246 bp in length
* 10750 10850: contig of 1001 bp in length
* 10851 11851: gap of unknown length
* 11852 13664: contig of 1714 bp in length
* 13665 13764: gap of unknown length
* 13765 15350: contig of 1586 bp in length
* 15351 17052: contig of 1602 bp in length
* 17053 17152: gap of unknown length
* 17153 18810: contig of 1658 bp in length
* 18811 18910: gap of unknown length
* 18911 20642: contig of 1732 bp in length
* 20643 20742: gap of unknown length
* 20743 22066: contig of 1324 bp in length
* 22067 22166: gap of unknown length
* 22167 24900: contig of 2734 bp in length
* 24901 25000: gap of unknown length
* 25001 28018: contig of 3018 bp in length
* 28019 28118: gap of unknown length
* 28119 30655: contig of 2537 bp in length
* 30656 30755: gap of unknown length
* 30756 33140: contig of 2385 bp in length
* 33141 33240: gap of unknown length
* 33241 35367: contig of 2127 bp in length
* 35368 35467: gap of unknown length
* 35468 37817: contig of 2350 bp in length
* 37818 37917: gap of unknown length
* 37918 41064: contig of 3147 bp in length
* 41065 41164: gap of unknown length
* 41165 44644: contig of 3480 bp in length
* 44645 44744: gap of unknown length

```

\* 47445 47442: contig of 2698 bp in length  
\* 47443 47542: gap of unknown length  
\* 47543 49411: contig of 1869 bp in length  
\* 49412 49511: gap of unknown length  
\* 49512 52288: contig of 2777 bp in length  
\* 52289 52388: gap of unknown length  
\* 52389 55771: contig of 3383 bp in length  
\* 55772 55871: gap of unknown length  
\* 55872 58117: contig of 2246 bp in length  
\* 58118 58217: gap of unknown length  
\* 58218 61622: contig of 3405 bp in length  
\* 61623 61722: gap of unknown length  
\* 61723 65588: contig of 3866 bp in length  
\* 65589 65688: gap of unknown length  
\* 65689 69895: contig of 4207 bp in length  
\* 69896 69995: gap of unknown length  
\* 69996 72794: contig of 2798 bp in length  
\* 72794 72893: gap of unknown length  
\* 72894 75602: contig of 2709 bp in length  
\* 75603 75703: gap of unknown length  
\* 75703 79032: contig of 3390 bp in length  
\* 79033 79192: gap of unknown length  
\* 79193 83380: contig of 4188 bp in length  
\* 83381 83480: gap of unknown length  
\* 83481 86437: contig of 3017 bp in length  
\* 86438 86597: gap of unknown length  
\* 86598 89889: contig of 3292 bp in length  
\* 89890 93787: gap of unknown length  
\* 93788 93887: contig of 3798 bp in length  
\* 93888 97050: contig of 3163 bp in length  
\* 97051 97150: gap of unknown length  
\* 97151 100721: contig of 3571 bp in length  
\* 100722 100821: gap of unknown length  
\* 100822 106757: contig of 5936 bp in length  
\* 106758 106857: gap of unknown length  
\* 106858 113422: contig of 6565 bp in length  
\* 113423 113523: gap of unknown length  
\* 113523 121377: contig of 7855 bp in length  
\* 121378 121477: gap of unknown length  
\* 121478 130444: contig of 8967 bp in length  
\* 130445 139345: gap of unknown length  
\* 139346 139445: contig of 8801 bp in length  
\* 139446 148719: gap of unknown length  
\* 148720 148819: contig of 9274 bp in length  
\* 148820 160872: contig of 12053 bp in length.

FEATURES

source  
1. .160872  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-745B12"

misc\_feature 1. .1335  
/note="assembly\_name:Contig12"  
misc\_feature 1436. .2709  
/note="assembly\_name:Contig13"  
misc\_feature 2810. .4388  
/note="assembly\_name:Contig15"  
misc\_feature 4489. .5988  
/note="assembly\_name:Contig16"  
misc\_feature 6089. .8110  
/note="assembly\_name:Contig17"  
misc\_feature 8211. .9403  
/note="assembly\_name:Contig18"  
misc\_feature 9504. .10749  
/note="assembly\_name:Contig19"  
misc\_feature 10850. .11850  
/note="assembly\_name:Contig2"  
misc\_feature 11951. .13664  
/note="assembly\_name:Contig20"  
misc\_feature 13765. .15350

misc\_feature /note="assembly\_name:Contig22"  
15451. .17052  
/note="assembly\_name:Contig23"  
17153. .18810  
/note="assembly\_name:Contig24"  
18911. .20642  
/note="assembly\_name:Contig25"  
20743. .22066  
/note="assembly\_name:Contig26"  
22167. .24900  
/note="assembly\_name:Contig27"  
25001. .28018  
/note="assembly\_name:Contig28"  
28119. .30655  
/note="assembly\_name:Contig29"  
30756. .33140  
/note="assembly\_name:Contig30"  
33241. .35367  
/note="assembly\_name:Contig31"  
35468. .37817  
/note="assembly\_name:Contig32"  
37918. .41064  
/note="assembly\_name:Contig33"  
41165. .44644  
/note="assembly\_name:Contig34"  
44745. .47442  
/note="assembly\_name:Contig35"  
47543. .49411  
/note="assembly\_name:Contig36"  
49512. .52288  
/note="assembly\_name:Contig37"

Query Match 84.0%; Score 16.8; DB 2; Length 160872;  
Best Local Similarity 90.0%; Pred. No. 5.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGCTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 153066 AGGGATTCTCTGAGTAGCAG 153047

Search completed: March 14, 2005, 20:42:21  
Job time : 1794 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 17:14:02 ; Search time 252 Seconds  
(without alignments)  
469.821 Million cell updates/sec

Title: US-10-672-399-10

Perfect score: 20

Sequence: 1 agggcgctctgtagtcagcag 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	1737	10 ADL13691	Adl13691 Osteoarth
C 2	20	100.0	1737	10 ADL13692	Adl13692 Osteoarth
C 3	20	100.0	2087	10 ADL13694	Adl13694 Osteoarth
C 4	20	100.0	2088	10 ADL13690	Adl13690 Osteoarth
C 5	20	100.0	2116	2 AAT99541	Aat99541 Human hya
C 6	20	100.0	2117	2 AAT96713	Aat96713 Human hya
C 7	20	100.0	231222	10 ADL13693	Adl13693 Osteoarth
C 8	17.4	87.0	1752	8 ABZ76734	Abz76734 Mouse hya
C 9	17.4	87.0	1752	10 AAD59442	Aad59442 Mouse hya
C 10	17.4	87.0	2102	2 AAT91655	Aat91655 Mouse hya
C 11	17.4	87.0	2102	2 AAZ10862	Aaz10862 Hyalurona
C 12	17.4	87.0	2102	3 AAZ88199	Aaz88199 Mouse hya
C 13	17.4	87.0	2102	3 AAA39987	Aaa39987 Murine HA
C 14	16.8	84.0	239	12 ACH89039	Ach89039 Human gen
C 15	16.8	84.0	278	3 AAC10010	Aac10010 Human sec
C 16	16.8	84.0	313	3 AAC02292	Aac02292 Human sec
C 17	16.8	84.0	595	12 ACH75339	Ach75339 Human gen
C 18	16.8	84.0	775	5 AAH19499	Aah19499 Human gal
C 19	16.8	84.0	80105	13 ABD33213	Abd33213 Murine ca
C 20	16.4	82.0	300	2 AAX98492	Aax98492 Human can

21	16.4	82.0	452	2 AAX99116	Aax99116 Human can
C 22	16.4	82.0	748	12 ADJ42940	Adj42940 Plant CDN
23	16.4	82.0	1611	10 ADB62940	Adb62940 Human CDN
24	16.4	82.0	1992	3 AAC77862	Aac77862 Human can
25	16.4	82.0	2057	4 AAH17632	Aah17632 Human CDN
26	16.4	82.0	2250	10 ADF83544	Adf83544 Human GRI
27	16.4	82.0	2508	13 ACN37249	Acn37249 Tumour-as
C 28	16.4	82.0	2665	6 ABL90233	Ab190233 Human pol
29	16.4	82.0	2753	4 AAH18559	Aah18559 Human CDN
30	16.4	82.0	2781	3 AAH72382	Aaa72382 Human nuc
C 31	16.4	82.0	3163	4 AAK52536	Aak52536 Human pol
32	16.4	82.0	3236	4 AAK51552	Aak51552 Human pol
33	16.4	82.0	3619	3 AAC75812	Aac75812 Human ORF
C 34	16.4	82.0	4560	12 ADQ23383	Adf83383 Human sof
35	16.4	82.0	15105	10 ADF83532	Adf83532 Human GRI
C 36	16.4	82.0	63411	12 ADQ97081	Adq97081 Mouse can
C 37	15.8	79.0	458	9 ACH41376	Ach41376 Human foe
38	15.8	79.0	525	4 AAH22902	Aah22902 Codon opt
39	15.8	79.0	525	10 ADK41220	Adk41220 DNA encod
40	15.8	79.0	531	6 ABS52517	Abse52517 DNA encod
C 41	15.8	79.0	596	8 ABZ53961	Abz53961 Aspergill
C 42	15.8	79.0	779	3 AAA01703	Aaa01703 Human col
C 43	15.8	79.0	882	11 ABD12260	Abd12260 Pseudomon
44	15.8	79.0	1203	11 ABD12216	Abd12216 Pseudomon
45	15.8	79.0	2069	13 ADR07772	Adr07772 Full leng
C 46	15.8	79.0	2614	2 AAV48219	Aav48219 Mouse str
C 47	15.8	79.0	2614	3 AAZ51045	Aaz51045 Mouse str
C 48	15.8	79.0	2928	12 ADQ97808	Adq97808 Mouse can
C 49	15.8	79.0	5000	4 AAS14505	Aas14505 Human GST
C 50	15.8	79.0	5727	10 ADK40991	Adk40991 Novel hum
C 51	15.8	79.0	5727	13 ADR15705	Adr15705 Kinase 41
C 52	15.8	79.0	5823	4 AAF22450	Aaf22450 Human bre
C 53	15.8	79.0	5866	8 ACC72055	Acc72055 BCU0518 g
C 54	15.8	79.0	6413	4 AAK52956	Aak52956 Human pol
C 55	15.8	79.0	6418	4 AAK51372	Aak51372 Human pol
C 56	15.8	79.0	6440	13 ACN39118	Acn39118 Tumour-as
C 57	15.8	79.0	47040	12 ADQ97807	Adq97807 Mouse can
C 58	15.8	79.0	53795	9 ADA02858	Ada02858 Human LMO
C 59	15.8	79.0	53795	10 ADB72596	Adb72596 Human LMO
C 60	15.8	79.0	53795	10 ADC85337	Adc85337 Mouse Lmo
C 61	15.8	79.0	53795	12 ADM74453	Adm74453 Human car
C 62	15.8	79.0	73465	6 ABQ88161	Abq88161 Human oet
C 63	15.8	79.0	110000	4 AAI99682_04	Continuation (5 of
64	15.8	79.0	110000	4 AAI99683_04	Continuation (5 of
65	15.8	79.0	114793	4 AAD08215_04	Continuation (5 of
66	15.8	79.0	153170	12 ADQ17382	Adq17382 Human sof
67	15.8	79.0	208765	12 ADQ97430	Adq97430 Mouse can
C 68	15.4	77.0	431	3 AAF08984	Aaf08984 Fusarium
C 69	15.4	77.0	433	8 ABX64635	Abx64635 Human gen
C 70	15.4	77.0	962	6 ABL39698	Ab139698 Human NS
C 71	15.4	77.0	1066	4 AAS22818	Aas22818 Human CDN
C 72	15.4	77.0	1066	13 ADS11519	Adsl11519 Human the
C 73	15.4	77.0	1255	4 AAS22582	Aas22582 Human CDN
C 74	15.4	77.0	1285	6 AAD36080	Aad36080 Human ATP
C 75	15.4	77.0	1285	6 AAD36081	Aad36081 Human ATP
C 76	15.4	77.0	2016	13 ACN37559	Acn37559 Tumour-as
C 77	15.4	77.0	2074	2 AAX04354	Aax04354 Human sec
C 78	15.4	77.0	15843	11 ACN44786	Acn44786 Human gen
C 79	15.2	76.0	21	13 ADS18343	Adsl8343 Murine TR
C 80	15.2	76.0	159	4 AAI55222	Aai55222 Probe #23
C 81	15.2	76.0	159	4 ABS48980	Abs48980 Human liv
C 82	15.2	76.0	380	12 ADP93531	Adp93531 Cotton ex
C 83	15.2	76.0	498	9 ACH36536	Ach36536 Human end
C 84	15.2	76.0	499	3 AAC00811	Aac00811 Human sec
C 85	15.2	76.0	503	9 ACH33753	Ach33753 Human end
C 86	15.2	76.0	548	5 AAS45111	Aas45111 cDNA enco
C 87	15.2	76.0	554	4 AA143036	Aa143036 Probe #11
C 88	15.2	76.0	554	4 ABS3680	Abas3680 Human liv
C 89	15.2	76.0	563	8 ABS53427	Abz53427 Aspergill
C 90	15.2	76.0	574	13 ADQ58168	Adq58168 Novel can
91	15.2	76.0	606	11 ABD02003	Abd02003 Pseudomon
92	15.2	76.0	684	10 ADJ32129	Adj32129 Human int
93	15.2	76.0	744	10 ADF58008	Adf58008 Human pol

C 94	15.2	76.0	834	6	AAD27736	Aad27736 Human ful	C 167	15.2	76.0	15515	8	AAL53548	Aal53548 Genomic D
C 95	15.2	76.0	895	6	AAD29597	Aad29597 Human cdn	C 168	15.2	76.0	25083	9	ADA02855	Ada02855 Mouse lmo
C 96	15.2	76.0	910	12	ADO35687	Ado35687 Novel mou	C 169	15.2	76.0	25083	10	ADB72593	Adb72593 Mouse lmo
C 97	15.2	76.0	960	4	ABLU08433	AbLU08433 Drosophill	C 170	15.2	76.0	25083	10	ADC85334	Adc85334 Human cbx
C 98	15.2	76.0	1077	10	ADE29368	Ade29368 Human eer	C 171	15.2	76.0	25083	12	ADM74450	Adm74450 Murine ca
C 99	15.2	76.0	1078	4	AAP83972	Aaf83972 Nucleotid	C 172	15.2	76.0	25231	4	AAL03112	Aal03112 Human rep
C 100	15.2	76.0	1101	12	ADH17419	Adh17419 Human NOV	C 173	15.2	76.0	49914	13	ABD33262	Abd33262 Murine ca
C 101	15.2	76.0	1110	12	ADH17423	Adh17423 Human NOV	C 174	15.2	76.0	63248	11	ACN45100	Acn45100 Mouse gen
C 102	15.2	76.0	1146	12	ADH17443	Adh17443 Human NOV	C 175	15.2	76.0	79640	13	ABD33007	Abd33007 Mouse can
C 103	15.2	76.0	1182	11	ABD01916	Abd01916 Pseudomon	C 176	15.2	76.0	112190	4	AHH44801	Aah44801 Human GPC
C 104	15.2	76.0	1203	12	ADH17425	Adh17425 Human NOV	C 177	15.2	76.0	120144	13	ABD33161	Abd33161 Murine ca
C 105	15.2	76.0	1203	12	ADH17421	Adh17421 Human NOV	C 178	15.2	76.0	129017	12	ADP84158	Adp84158 Human AST
C 106	15.2	76.0	1203	12	ADH17427	Adh17427 Human NOV	C 179	15.2	76.0	150085	13	ABD32809	Abd32809 Mouse can
C 107	15.2	76.0	1218	12	ADH17435	Adh17435 Human NOV	C 180	15.2	76.0	173810	6	ABN85752	Abn85752 Mouse chr
C 108	15.2	76.0	1218	12	ADH17447	Adh17447 Human NOV	C 181	15.2	76.0	247509	10	ADL13684	Adl13684 Osteoarth
C 109	15.2	76.0	1230	6	AAD27743	Aad27743 Human tra	C 182	15.2	75.0	456	9	ACH33386	Ach33386 Human end
C 110	15.2	76.0	1282	10	ADE79051	Ade79051 Human pro	C 183	15.2	75.0	552	12	ACH78684	Ach78684 Human gen
C 111	15.2	76.0	1299	11	ABD01892	Abd01892 Pseudomon	C 184	15.2	75.0	1379	4	AAS22470	Aas22470 Human cdn
C 112	15.2	76.0	1314	4	AAP83973	Aaf83973 Nucleotid	C 185	15.2	75.0	10622	4	AAK68203	Aak68203 Human imm
C 113	15.2	76.0	1314	4	AAP83971	Aaf83971 Human SER	C 186	15.2	75.0	11876	12	ADQ18981	Adq18981 Human sof
C 114	15.2	76.0	1314	10	ADE29366	Ade29366 Human cdn	C 187	15.2	75.0	11919	12	ADQ23338	Adq23338 Human sof
C 115	15.2	76.0	1314	10	ADE29369	Ade29369 Human eer	C 188	15.2	75.0	56743	4	AAK68202	Aak68202 Human imm
C 116	15.2	76.0	1314	12	ADH17441	Adh17441 Human NOV	C 189	15.2	75.0	56743	4	AAK81760	Aak81760 Human imm
C 117	15.2	76.0	1341	8	ABZ22865	Abz22865 Human den	C 190	15.2	75.0	65464	8	ABX13172	Abx13172 Human gen
C 118	15.2	76.0	1407	10	ADE79039	Ade79039 Human pro	C 191	15.2	75.0	200418	11	ACN44226	Acn44226 Human gen
C 119	15.2	76.0	1409	10	ADH79066	Ade79066 Human pro	C 192	14.8	74.0	295	2	AAZ27708	Aaz27708 Human DNA
C 120	15.2	76.0	1434	8	ABZ22864	Abz22864 Human den	C 193	14.8	74.0	411	9	ACH17071	Ach17071 Human adu
C 121	15.2	76.0	1448	10	ADE79040	Ade79040 Human pro	C 194	14.8	74.0	455	9	ACH40376	Ach40376 Human foe
C 122	15.2	76.0	1568	8	ACC46449	Acc46449 Human dit	C 195	14.8	74.0	467	4	AAI11592	Aai11592 Probe #15
C 123	15.2	76.0	1574	6	ADA30575	Ada30575 Human pro	C 196	14.8	74.0	467	4	ABA53281	Abas3281 Human foe
C 124	15.2	76.0	1614	6	ABK31788	Abk31788 DNA encod	C 197	14.8	74.0	467	4	AAI32884	Aai32884 Probe #15
C 125	15.2	76.0	1626	12	ADH17431	Adh17431 Human NOV	C 198	14.8	74.0	467	4	ABA42858	Abas42858 Human bre
C 126	15.2	76.0	1689	4	AAD05796	Aad05796 Human tra	C 199	14.8	74.0	467	4	ABA23055	Abas23055 Probe #15
C 127	15.2	76.0	1689	10	ADH10392	Adh10392 Human cel	C 200	14.8	74.0	467	4	AAK26984	Aak26984 Human bon
C 128	15.2	76.0	1689	12	ADJ46916	Adj46916 Human tra	C 201	14.8	74.0	467	4	AAK01539	Aak01539 Human bra
C 129	15.2	76.0	1707	12	ADH17439	Adh17439 Human NOV	C 202	14.8	74.0	467	4	ABS26572	Abs26572 Human liv
C 130	15.2	76.0	1708	12	ADH17457	Adh17457 Human NOV	C 203	14.8	74.0	467	5	AAI01516	Aai01516 Probe #15
C 131	15.2	76.0	1748	6	AAD27734	Aad27734 Human tra	C 204	14.8	74.0	467	6	ABS01571	Abs01571 Human gen
C 132	15.2	76.0	1771	12	ADH17433	Adh17433 Human NOV	C 205	14.8	74.0	491	3	AAO07469	Aao07469 Human sec
C 133	15.2	76.0	1779	12	ADH17429	Adh17429 Human NOV	C 206	14.8	74.0	513	6	ABN19118	Abn19118 Human ORF
C 134	15.2	76.0	1858	2	AAQ56202	Aaq56202 Partial p	C 207	14.8	74.0	552	10	ADD29823	Add29823 Human tum
C 135	15.2	76.0	1858	2	AAQ30126	Aaq30126 Plasmid p	C 208	14.8	74.0	572	4	AAI20806	Aai20806 Probe #10
C 136	15.2	76.0	1858	2	NAT38418	Nat38418 pCK1 cDNA	C 209	14.8	74.0	572	4	ABA65874	Abas65874 Human foe
C 137	15.2	76.0	1899	4	ABLI2039	Abli2039 Drosophill	C 210	14.8	74.0	572	4	AAI46037	Aai46037 Probe #14
C 138	15.2	76.0	2057	4	AAD05797	Aad05797 Human tra	C 211	14.8	74.0	572	4	ABA47987	Abas47987 Human bre
C 139	15.2	76.0	2067	10	ADH10394	Adh10394 Human cel	C 212	14.8	74.0	572	4	ABA32960	Abas32960 Probe #11
C 140	15.2	76.0	2067	12	ADJ46918	Adj46918 Human tra	C 213	14.8	74.0	572	4	AAK40017	Aak40017 Human bon
C 141	15.2	76.0	2192	4	AAH14850	Aah14850 Human cdn	C 214	14.8	74.0	572	4	AAK14286	Aak14286 Human bra
C 142	15.2	76.0	2237	4	ABL23356	AbL23356 Drosophill	C 215	14.8	74.0	572	4	ABS39603	Abs39603 Human liv
C 143	15.2	76.0	2310	12	ADH17437	Adh17437 Human NOV	C 216	14.8	74.0	572	5	AAI06514	Aai06514 Probe #65
C 144	15.2	76.0	2360	10	ADE79041	Ade79041 Human pro	C 217	14.8	74.0	572	6	ABS14103	Abs14103 Human gen
C 145	15.2	76.0	2393	10	ADH1742	Ade311742 Human 291	C 218	14.8	74.0	660	12	ADL81735	Adl81735 P. aerugi
C 146	15.2	76.0	2410	6	ABK12897	Abk12897 Human pro	C 219	14.8	74.0	707	12	ADJ42941	Adj42941 Plant cdn
C 147	15.2	76.0	2432	12	ADH17453	Adh17453 Human NOV	C 220	14.8	74.0	723	3	AAF12135	Aaf12135 Aspergill
C 148	15.2	76.0	2432	12	ADH17451	Adh17451 Human NOV	C 221	14.8	74.0	774	11	ABD11205	Abd11205 Pseudomon
C 149	15.2	76.0	2432	12	ADH17449	Adh17449 Human NOV	C 222	14.8	74.0	785	12	ADJ43613	Adj43613 Plant cdn
C 150	15.2	76.0	2432	12	ADH17413	Adh17413 Human NOV	C 223	14.8	74.0	878	4	AAK93648	Aak93648 Human cdn
C 151	15.2	76.0	2643	3	AAA46557	Aaa46557 cDNA sequ	C 224	14.8	74.0	878	4	AAK91982	Aak91982 Human cdn
C 152	15.2	76.0	2643	6	ABK72680	Abk72680 DNA encod	C 225	14.8	74.0	878	12	ADL30075	Adl30075 3' end of
C 153	15.2	76.0	2682	13	ADR10444	Adr10444 Full leng	C 226	14.8	74.0	878	12	ADL28409	Adl28409 5' end of
C 154	15.2	76.0	2682	5	AAS44923	Aas44923 cDNA enco	C 227	14.8	74.0	884	2	AXX99070	Aax99070 Human val
C 155	15.2	76.0	2865	4	ABL04669	AbL04669 Drosophill	C 228	14.8	74.0	1002	8	ADA69388	Ada69388 Rice gene
C 156	15.2	76.0	3103	13	ACN42023	Acn42023 Human dia	C 229	14.8	74.0	1037	4	ABL10383	AbL10383 Drosophill
C 157	15.2	76.0	3244	4	ABL08432	AbL08432 Drosophill	C 230	14.8	74.0	1089	11	ABD11386	Abd11386 Pseudomon
C 158	15.2	76.0	3472	4	ABL07626	AbL07626 Drosophill	C 231	14.8	74.0	1158	3	AZ292421	Aaz292421 DNA encod
C 159	15.2	76.0	4001	10	ADF18694	Adf18694 Human reg	C 232	14.8	74.0	1158	10	ACA62465	Acas62465 DNA encod
C 160	15.2	76.0	4219	4	ABL04902	AbL04902 Drosophill	C 233	14.8	74.0	1168	6	ABS65599	Abs65599 Mouse gen
C 161	15.2	76.0	5110	6	ABL2038	AbL2038 Drosophill	C 234	14.8	74.0	1254	2	AAT43482	Aat43482 ATM gene
C 162	15.2	76.0	5728	6	ABK2572	Abk2572 Rat sequ	C 235	14.8	74.0	1254	4	AAF57634	Aaf57634 ATM genom
C 163	15.2	76.0	5775	10	ADB58556	AdB58556 Toxicity-	C 236	14.8	74.0	1389	5	AAS87543	Aas87543 DNA encod
C 164	15.2	76.0	5775	10	ADB53170	AdB53170 Primary r	C 237	14.8	74.0	1465	5	AAS91285	Aas91285 DNA encod
C 165	15.2	76.0	6164	4	ABL04668	AbL04668 Drosophill	C 238	14.8	74.0	1511	5	AAS89099	Aas89099 DNA encod
C 166	15.2	76.0	7513	10	ADC68982	Adc68982 Human GPC	C 239	14.8	74.0	1537	6	ABV77431	Abv77431 Human syn

c 240	14.8	74.0	1560	2	AAQ54050	AaQ54050 Human bon	313	14.8	74.0	9169	2	AAT68760	Aat68760 ATM mutan
c 241	14.8	74.0	1640	2	AAT45868	Aat45868 Human bon	314	14.8	74.0	9169	2	AAT73811	Aat73811 ATM gene
c 242	14.8	74.0	2000	10	ACC60866	Acc60866 Gene sequ	315	14.8	74.0	9169	2	AAT73808	Aat73808 ATM gene
c 243	14.8	74.0	2000	10	ADK62245	AdK62245 Disease t	316	14.8	74.0	9170	2	AAT68762	Aat68762 ATM mutan
c 244	14.8	74.0	2007	5	AAS88300	Aas88300 DNA encod	317	14.8	74.0	9171	2	AAT68764	Aat68764 ATM mutan
c 245	14.8	74.0	2245	8	ACC79893	Acc79893 Human DEL	318	14.8	74.0	9171	2	AAT68765	Aat68765 ATM mutan
c 246	14.8	74.0	2271	11	ADJ11513	AdJ11513 Rice DNA	319	14.8	74.0	9171	2	AAT68773	Aat68773 ATM mutan
c 247	14.8	74.0	2484	11	ADM02858	AdM02858 Human CDN	320	14.8	74.0	9171	2	AAT68781	Aat68781 ATM mutan
c 248	14.8	74.0	2532	10	ADN39337	AdN39337 Human gen	321	14.8	74.0	9171	2	AAT68768	Aat68768 ATM mutan
c 249	14.8	74.0	2674	11	ADN39337	AdN39337 Cancer/an	322	14.8	74.0	9171	2	AAT68782	Aat68782 ATM mutan
c 250	14.8	74.0	2691	12	ADQ23720	AdQ23720 Human sof	323	14.8	74.0	9171	2	AAT43497	Aat43497 ATM open
c 251	14.8	74.0	2971	10	ADE15656	AdE15656 Human str	324	14.8	74.0	9171	2	AAT68733	Aat68733 ATM mutan
c 252	14.8	74.0	3012	11	ADM03028	AdM03028 Human CDN	325	14.8	74.0	9171	2	AAT68733	Aat68733 ATM mutan
c 253	14.8	74.0	3142	10	ADE56315	AdE56315 Human gen	326	14.8	74.0	9171	2	AAT45419	Aat45419 Ataxia-te
c 254	14.8	74.0	3142	10	ADE56319	AdE56319 Human gen	327	14.8	74.0	9171	2	AAT73810	Aat73810 ATM gene
c 255	14.8	74.0	3142	10	ADD46110	AdD46110 Human gen	328	14.8	74.0	9171	4	AAF57591	Aaf57591 ATM gene
c 256	14.8	74.0	3238	4	AAK94339	AaK94339 Human ful	329	14.8	74.0	9171	4	AAD21841	Aad21841 Human ATM
c 257	14.8	74.0	3238	12	ADL31002	AdL31002 Full leng	330	14.8	74.0	9171	4	AAD21843	Aad21843 Human ATM
c 258	14.8	74.0	3408	8	AAD52899	AdA52899 Mouse twe	331	14.8	74.0	9171	4	AAD21871	Aad21871 Human ATM
c 259	14.8	74.0	3495	13	ACN41849	AcN41849 Human dia	332	14.8	74.0	9171	4	AAD21836	Aad21836 Human ATM
c 260	14.8	74.0	3589	8	ABT32175	AbT32175 Human neu	333	14.8	74.0	9171	4	AAD21846	Aad21846 Human ATM
c 261	14.8	74.0	5080	8	ACD13351	AcD13351 Human DNA	334	14.8	74.0	9171	4	AAD21840	Aad21840 Human ATM
c 262	14.8	74.0	5912	2	AT4743519	Aat4743519 ATM gene	335	14.8	74.0	9171	4	AAD21848	Aad21848 Human ATM
c 263	14.8	74.0	5912	2	AT4743519	Aat4743519 ATM gene	336	14.8	74.0	9171	4	AAD21853	Aad21853 Human ATM
c 264	14.8	74.0	5912	4	AAF57590	Aaf57590 Nucleotid	337	14.8	74.0	9171	4	AAD21837	Aad21837 Human ATM
c 265	14.8	74.0	6487	8	ABZ74409	AbZ74409 Secreted	338	14.8	74.0	9171	4	AAD21845	Aad21845 Human ATM
c 266	14.8	74.0	6487	8	ADA98861	AdA98861 Human sec	339	14.8	74.0	9171	4	AAD21845	Aad21845 Human ATM
c 267	14.8	74.0	6487	8	ADA44475	AdA44475 Human sec	340	14.8	74.0	9171	4	AAD21847	Aad21847 Human ATM
c 268	14.8	74.0	6487	10	ABZ67966	AbZ67966 Human sec	341	14.8	74.0	9171	4	AAD21851	Aad21851 Human ATM
c 269	14.8	74.0	7207	6	ABU54575	AbU54575 Mouse per	342	14.8	74.0	9171	4	AAD21842	Aad21842 Human ATM
c 270	14.8	74.0	8341	2	AAT68772	Aat68772 ATM mutan	343	14.8	74.0	9171	4	AAD21850	Aad21850 Human ATM
c 271	14.8	74.0	8382	4	AAK70731	AaK70731 Human imm	344	14.8	74.0	9171	4	AAD21849	Aad21849 Human ATM
c 272	14.8	74.0	8682	4	ABL10382	AbL10382 Drosophil	345	14.8	74.0	9171	4	AAD21856	Aad21856 Human ATM
c 273	14.8	74.0	8712	5	AH24106	AaH24106 Mouse Per	346	14.8	74.0	9171	4	AAD21834	Aad21834 Human ATM
c 274	14.8	74.0	8767	2	AAT68735	Aat68735 ATM mutan	347	14.8	74.0	9171	4	AAD21852	Aad21852 Human ATM
c 275	14.8	74.0	8768	2	AAT73812	Aat73812 ATM gene	348	14.8	74.0	9171	4	AAD21832	Aad21832 Human ATM
c 276	14.8	74.0	8799	2	AAT68758	Aat68758 ATM mutan	349	14.8	74.0	9171	4	AAD21872	Aad21872 Human ATM
c 277	14.8	74.0	8816	2	AAT68740	Aat68740 ATM mutan	350	14.8	74.0	9171	4	AAD21855	Aad21855 Human ATM
c 278	14.8	74.0	8873	2	AAT68750	Aat68750 ATM mutan	351	14.8	74.0	9171	4	AAD21831	Aad21831 Human ATM
c 279	14.8	74.0	8970	2	AAT68771	Aat68771 ATM mutan	352	14.8	74.0	9171	4	AAD21835	Aad21835 Human ATM
c 280	14.8	74.0	8996	2	AAT68754	Aat68754 ATM mutan	353	14.8	74.0	9171	4	AAD21844	Aad21844 Human ATM
c 281	14.8	74.0	8996	2	AAT68783	Aat68783 ATM mutan	354	14.8	74.0	9171	4	AAD21854	Aad21854 Human ATM
c 282	14.8	74.0	8997	2	AAT68756	Aat68756 ATM mutan	355	14.8	74.0	9171	4	AAD21780	Aad21780 Human ATM
c 283	14.8	74.0	8997	2	AAT68774	Aat68774 ATM mutan	356	14.8	74.0	9171	4	AAD21833	Aad21833 Human ATM
c 284	14.8	74.0	9006	2	AAT68775	Aat68775 ATM mutan	357	14.8	74.0	9171	4	AAD21839	Aad21839 Human ATM
c 285	14.8	74.0	9006	2	AAT68753	Aat68753 ATM mutan	358	14.8	74.0	9171	10	ADH68996	Adh68996 ATM gene
c 286	14.8	74.0	9012	2	AAT68741	Aat68741 ATM mutan	359	14.8	74.0	9171	10	ADH69001	Adh69001 ATM gene
c 287	14.8	74.0	9018	2	AAT68736	Aat68736 ATM mutan	360	14.8	74.0	9171	10	ADH68992	Adh68992 ATM gene
c 288	14.8	74.0	9022	2	AAT68745	Aat68745 ATM mutan	361	14.8	74.0	9171	10	ADH68997	Adh68997 ATM gene
c 289	14.8	74.0	9029	2	AAT68739	Aat68739 ATM mutan	362	14.8	74.0	9171	10	ADH68999	Adh68999 ATM gene
c 290	14.8	74.0	9032	2	AAT68752	Aat68752 ATM mutan	363	14.8	74.0	9171	10	ADH68989	Adh68989 ATM gene
c 291	14.8	74.0	9044	2	AAT68755	Aat68755 ATM mutan	364	14.8	74.0	9171	10	ADH68988	Adh68988 ATM gene
c 292	14.8	74.0	9046	2	AAT68770	Aat68770 ATM mutan	365	14.8	74.0	9171	10	ADH69004	Adh69004 Human ATM
c 293	14.8	74.0	9046	2	AAT68761	Aat68761 ATM mutan	366	14.8	74.0	9171	10	ADH69011	Adh69011 Human ATM
c 294	14.8	74.0	9066	2	AAT68744	Aat68744 ATM mutan	367	14.8	74.0	9171	10	ADH69013	Adh69013 Human ATM
c 295	14.8	74.0	9081	2	AAT68759	Aat68759 ATM mutan	368	14.8	74.0	9171	10	ADH69007	Adh69007 Human ATM
c 296	14.8	74.0	9088	2	AAT68757	Aat68757 ATM mutan	369	14.8	74.0	9171	10	ADH69008	Adh69008 Human ATM
c 297	14.8	74.0	9160	2	AAT68749	Aat68749 ATM mutan	370	14.8	74.0	9171	10	ADH69010	Adh69010 Human ATM
c 298	14.8	74.0	9162	2	AAT68777	Aat68777 ATM mutan	371	14.8	74.0	9171	10	ADH69012	Adh69012 Human ATM
c 299	14.8	74.0	9164	2	AAT68751	Aat68751 ATM mutan	372	14.8	74.0	9171	10	ADH68990	Adh68990 ATM gene
c 300	14.8	74.0	9165	2	AAT68785	Aat68785 ATM mutan	373	14.8	74.0	9171	10	ADH68993	Adh68993 ATM gene
c 301	14.8	74.0	9165	2	AAT68778	Aat68778 ATM mutan	374	14.8	74.0	9171	10	ADH68995	Adh68995 ATM gene
c 302	14.8	74.0	9166	2	AAT68746	Aat68746 ATM mutan	375	14.8	74.0	9171	10	ADH68906	Adh68906 Human ATM
c 303	14.8	74.0	9166	2	AAT68738	Aat68738 ATM mutan	376	14.8	74.0	9171	10	ADH68932	Adh68932 Human ATM
c 304	14.8	74.0	9166	2	AAT68743	Aat68743 ATM mutan	377	14.8	74.0	9171	10	ADH69000	Adh69000 Human ATM
c 305	14.8	74.0	9167	2	AAT68742	Aat68742 ATM mutan	378	14.8	74.0	9171	10	ADH68986	Adh68986 Human ATM
c 306	14.8	74.0	9167	2	AAT68769	Aat68769 ATM mutan	379	14.8	74.0	9171	10	ADH68998	Adh68998 Human ATM
c 307	14.8	74.0	9168	2	AAT68776	Aat68776 ATM mutan	380	14.8	74.0	9171	10	ADH69003	Adh69003 ATM gene
c 308	14.8	74.0	9168	2	AAT68780	Aat68780 ATM mutan	381	14.8	74.0	9171	10	ADH69002	Adh69002 ATM gene
c 309	14.8	74.0	9169	2	AAT68731	Aat68731 ATM mutan	382	14.8	74.0	9171	10	ADH69005	Adh69005 Human ATM
c 310	14.8	74.0	9169	2	AAT68763	Aat68763 ATM mutan	383	14.8	74.0	9171	10	ADH69009	Adh69009 Human ATM
c 311	14.8	74.0	9169	2	AAT68734	Aat68734 ATM mutan	384	14.8	74.0	9171	10	ADH68987	Adh68987 Human ATM
c 312	14.8	74.0	9169	2	AAT68766	Aat68766 ATM mutan	385	14.8	74.0	9171	10	ADH68991	Adh68991 ATM gene



PI Jones KA, Schafer A;  
 XX WPI; 2003-559141/52.  
 XX Determining susceptibility of an individual to joint space narrowing,  
 PT osteophyte development and/or joint pain comprises identifying whether  
 PT the individual has at least one polymorphism in a polynucleotide encoding  
 PT a protein.  
 XX Disclosure; SEQ ID NO 223; 297pp; English.  
 XX The invention relates to a method of determining susceptibility of an  
 CC individual to joint space narrowing and/or osteophyte development and/or  
 CC joint pain comprising identifying whether the individual has at least one  
 CC polymorphism in a polynucleotide encoding at least one of the protein  
 CC listed in the specification. The methods, composition and agent are  
 CC useful for modulating the susceptibility of an individual to joint space  
 CC narrowing and/or osteophyte development and/or joint pain that is  
 CC associated with a disease, preferably osteoarthritis. The cell line and  
 CC the non-human animal are useful for screening for an agent for diagnosing  
 CC an individual having susceptibility to joint space narrowing and/or  
 CC osteophyte development and/or joint pain. This sequence corresponds to  
 CC the polynucleotide encoding a protein listed in the specification. (Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences).  
 XX Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 20; DB 10; Length 1737;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 DB 1099 AGGGCGTCTCTGAGTAGCAG 1080  
 RESULT 2  
 ADL13692/c  
 ID ADL13692 standard; DNA; 1737 BP.  
 XX AC ADL13692;  
 XX DT 06-MAY-2004 (first entry)  
 XX DE Osteoarthritis-associated polymorphic nucleotide #224.  
 XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
 KW joint space narrowing; osteophyte development; joint pain;  
 KW osteoarthritis; SNP; single nucleotide polymorphism.  
 XX OS Homo sapiens.  
 XX PN WO2003054166-A2.  
 XX PD 03-JUL-2003.  
 XX PF 19-DEC-2002; 2002WO-US041225.  
 XX PR 20-DEC-2001; 2001US-0342603P.  
 XX PA (INCY-) INCYTE GENOMICS INC.  
 XX PI Jones KA, Schafer A;  
 XX WPI; 2003-559141/52.  
 XX Determining susceptibility of an individual to joint space narrowing,  
 PT osteophyte development and/or joint pain comprises identifying whether  
 PT the individual has at least one polymorphism in a polynucleotide encoding  
 PT a protein.

PS Disclosure; SEQ ID NO 224; 297pp; English.  
 XX The invention relates to a method of determining susceptibility of an  
 CC individual to joint space narrowing and/or osteophyte development and/or  
 CC joint pain comprising identifying whether the individual has at least one  
 CC polymorphism in a polynucleotide encoding at least one of the protein  
 CC listed in the specification. The methods, composition and agent are  
 CC useful for modulating the susceptibility of an individual to joint space  
 CC narrowing and/or osteophyte development and/or joint pain that is  
 CC associated with a disease, preferably osteoarthritis. The cell line and  
 CC the non-human animal are useful for screening for an agent for diagnosing  
 CC an individual having susceptibility to joint space narrowing and/or  
 CC osteophyte development and/or joint pain. This sequence corresponds to  
 CC the polynucleotide encoding a protein listed in the specification. (Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences).  
 XX Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 20; DB 10; Length 1737;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 DB 1099 AGGGCGTCTCTGAGTAGCAG 1080  
 RESULT 3  
 ADL13694/c  
 ID ADL13694 standard; DNA; 2087 BP.  
 XX AC ADL13694;  
 XX DT 06-MAY-2004 (first entry)  
 XX DE Osteoarthritis-associated polymorphic nucleotide #226.  
 XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
 KW joint space narrowing; osteophyte development; joint pain;  
 KW osteoarthritis; SNP; single nucleotide polymorphism.  
 XX OS Homo sapiens.  
 XX PN WO2003054166-A2.  
 XX PD 03-JUL-2003.  
 XX PF 19-DEC-2002; 2002WO-US041225.  
 XX PR 20-DEC-2001; 2001US-0342603P.  
 XX PA (INCY-) INCYTE GENOMICS INC.  
 XX PI Jones KA, Schafer A;  
 XX WPI; 2003-559141/52.  
 XX Determining susceptibility of an individual to joint space narrowing,  
 PT osteophyte development and/or joint pain comprises identifying whether  
 PT the individual has at least one polymorphism in a polynucleotide encoding  
 PT a protein.  
 XX Disclosure; SEQ ID NO 226; 297pp; English.  
 XX The invention relates to a method of determining susceptibility of an  
 CC individual to joint space narrowing and/or osteophyte development and/or  
 CC joint pain comprising identifying whether the individual has at least one  
 CC polymorphism in a polynucleotide encoding at least one of the protein  
 CC listed in the specification. The methods, composition and agent are  
 CC useful for modulating the susceptibility of an individual to joint space  
 CC narrowing and/or osteophyte development and/or joint pain that is

CC associated with a disease, preferably osteoarthritis. The cell line and  
CC the non-human animal are useful for screening for an agent for diagnosing  
CC an individual having susceptibility to joint space narrowing and/or  
CC osteophyte development and/or joint pain. This sequence corresponds to  
CC the polynucleotide encoding a protein listed in the specification. (Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences).

XX SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2087;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
Db 1134 AGGGCGTCTCTGAGTAGCAG 1115

RESULT 4  
ADL13690/c  
ID ADL13690 standard; DNA; 2088 BP.

XX AC ADL13690;

XX DT 06-MAY-2004 (first entry)

XX DE Osteoarthritis-associated polymorphic nucleotide #222.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

XX joint space narrowing; osteophyte development; joint pain;

XX osteoarthritis; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX WO2003054166-A2.

XX PD 03-JUL-2003.

XX PF 19-DEC-2002; 2002WO-US041225.

XX PR 20-DEC-2001; 2001US-0342603P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,  
PT osteophyte development and/or joint pain comprises identifying whether  
PT the individual has at least one polymorphism in a polynucleotide encoding  
PT a protein.

PS Disclosure; SEQ ID NO 222; 297pp; English.

XX The invention relates to a method of determining susceptibility of an  
CC individual to joint space narrowing and/or osteophyte development and/or  
CC joint pain comprising identifying whether the individual has at least one  
CC polymorphism in a polynucleotide encoding at least one of the protein  
CC listed in the specification. The methods, composition and agent are  
CC useful for modulating the susceptibility of an individual to joint space  
CC narrowing and/or osteophyte development and/or joint pain that is  
CC associated with a disease, preferably osteoarthritis. The cell line and  
CC the non-human animal are useful for screening for an agent for diagnosing  
CC an individual having susceptibility to joint space narrowing and/or  
CC osteophyte development and/or joint pain. This sequence corresponds to  
CC the polynucleotide encoding a protein listed in the specification. (Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences).

SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2088;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
Db 1134 AGGGCGTCTCTGAGTAGCAG 1115

RESULT 5  
AAT99541/c

ID AAT99541 standard; cDNA; 2116 BP.

XX AC AAT99541;

XX DT 21-MAY-1998 (first entry)

XX DE Human hyaluronan synthase cDNA clone 30C.

XX KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;  
XX wound healing; vulnarary; tissue repair; scar; keloid; therapy; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 36..1172

XX FT polyA\_signal /\*tag= a

XX FT 2066..2071 /\*tag= b

XX PN WO9740174-A1.

XX PD 30-OCT-1997.

XX PF 17-APR-1997; 97WO-US006350.

XX PR 22-APR-1996; 96US-00635552.

XX PA (LEUK-) LEUKOSITE INC.

XX PI Briskin MJ;

XX WPI; 1997-549359/50.

XX P-PSDB; AAW26765.

XX Human hyaluronan synthase - useful for recombinant production of  
PT hyaluronic acid for wound healing, tissue repair and reducing  
PT hypertrophic scar and keloid formation.

PS Claim 3; Page 36-38; 58pp; English.

XX cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)  
CC (see AAW26765), an enzyme involved in the synthesis of hyaluronan  
CC (hyaluronic acid) and which has the ability to confer cell adhesion by  
CC the lymphocyte receptor CD44. Clone 30C was isolated using an expression  
CC cloning system developed to isolate cDNA clones that encode proteins that  
CC confer adhesion of the murine T cell lymphoma TK1. A human mesenteric  
CC lymph node expression library was constructed that, upon transfection  
CC into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells  
CC to some of the transfectants. The isolated clone can be utilised in a  
CC claimed method for producing HAS in host cells. Such host cells are used  
CC in a claimed method for the production of hyaluronan. Hyaluronan is  
CC useful for wound healing and tissue repair, and can reduce or prevent  
CC hypertrophic scars and keloid formation. It is also used in eye surgery  
CC as a replacement for vitreous fluid

XX SQ Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
QY 1 AGGCGCTCTCTGAGTAGCAG 20
DB 1134 AGGCGCTCTCTGAGTAGCAG 1115

RESULT 6
AAT96713/c
ID AAT96713 standard; DNA; 2117 BP.
XX
AC AAT96713;
XX
XX
DT 22-APR-1998 (first entry)
DE
DE Human hyaluronate synthetase coding sequence.
KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
KW cosmetic preparation; gene therapy; carcinogenesis; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 149..1780
FT /*tag= a
XX
PN WO9738113-A1.
XX
XX 16-OCT-1997.
XX
XX 31-MAR-1997; 97WO-JP001111.
XX
XX 05-APR-1996; 96JP-00084326.
XX 30-APR-1996; 96JP-00109663.
XX
XX (SEKG ) SEIKAGAKU CORP.
XX
XX Itano N, Kimata K;
XX
XX WPI; 1997-512726/47.
XX P-PSDB; AAW36503.
XX
XX DNA encoding human hyaluronate synthetase - for industrial scale
XX production of hyaluronic acid used in generating anti-carcinogenic drugs
XX or for cosmetics.
XX
XX Claim 3; Page 23-27; 35pp; Japanese.
XX
XX This sequence encodes a human hyaluronate synthetase, and is the coding
XX sequence of the invention. The encoded enzyme is useful for industrial
XX scale production of hyaluronic acid for use in the preparation of drugs
XX and cosmetics. The drugs can also be used in compositions for the
XX treatment of disorders involving the lowering of hyaluronic acid
XX production. The peptides may be used for the preparation of antibodies
XX recognising hyaluronate synthetase, e.g. for diagnostic purposes.
XX Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
XX treatment of carcinogenesis
XX
XX Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 20; DB 2; Length 2117;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20
DB 1142 AGGCGCTCTCTGAGTAGCAG 1123

RESULT 7
ADL13693
ID ADL13693 standard; DNA; 231222 BP.
XX
XX ADL13693;
AC
```

```
XX
DT 06-MAY-2004 (first entry)
DE
DE Osteoarthritis-associated polymorphic nucleotide #225.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
XX WO2003054166-A2.
XX
XX 03-JUL-2003.
XX
XX 19-DEC-2002; 2002WO-US041225.
XX
XX 20-DEC-2001; 2001US-0342603P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 225; 297pp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;
XX
Query Match 100.0%; Score 20; DB 10; Length 231222;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20
DB 143286 AGGCGCTCTCTGAGTAGCAG 143305

RESULT 8
ABZ76734/c
ID ABZ76734 standard; cDNA; 1752 BP.
XX
XX ABZ76734;
XX
XX 01-APR-2000; (first entry)
XX
XX Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.
XX
XX Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
KW osteoarthritis; gene; ss.
```

```

XX OS Mus sp.
XX FH Key Location/Qualifiers
XX FT CDS 1..1752
XX FT /*tag= a
XX FT /product= "hyaluronan synthase 1 (HAS1)"
XX PN WO2003006068-A1.
XX PD 23-JAN-2003.
XX PF 10-JUL-2001; 2001WO-US021785.
XX XX 10-JUL-2001; 2001WO-US021785.
XX PR (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
XX PA Dehazya P, Chen W;
XX PI WPI; 2003-221664/21.
XX DR P-PSDB; ABP96028.
XX XX Novel dihydrazone derivatized hyaluronic acid/nucleic acid bioconjugate
PT for treating dry eye syndrome, has derivatized hyaluronic acid
PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
PT activity.
XX XX Claim 19; Page 56; 62pp; English.
XX PS The present invention describes a dihydrazone derivatized hyaluronic acid
CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatized HA linked to NA
CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
CC antiarthritic activities, and can be used in gene therapy and as an
CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
CC human corneal epithelial cell. (I) is useful for transfecting a cell of
CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
CC (I) is useful for treating dry eye syndrome in an individual. (I) is
CC useful in gene therapy applications for the treatment of a variety of
CC medical conditions including dry eye syndrome or other medical conditions
CC where an increase in the production of (HA) in the eye would be
CC therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
CC useful for inhibiting angiogenesis for the treatment of macular
CC degeneration or genes related to lipid biosynthesis that helps to restore
CC the lipid component of the tear film, and as reagents for in vitro
CC transformation of any cell, preferably a eukaryotic cell, more preferably
CC a human eye cell. The present sequence encodes mouse HAS1 which is used
CC in the exemplification of the present invention
XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 87.0%; Score 17.4; DB 8; Length 1752;
Best Local Similarity 94.7%; Pred. No. 91;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGCTCTCGAGTAGCA 19
Db 1114 AGGGCGCTCTCGAGTAGCA 1096

RESULT 9
AAD59442/C
ID AAD59442 standard; DNA; 1752 BP.
XX AC AAD59442;
XX AC AAD59442;
XX DT 18-DEC-2003 (first entry)
XX XX

RESULT 10
AAT91655/C
ID AAT91655 standard; DNA; 2102 BP.
XX AC AAT91655;
XX AC AAT91655;
XX DT 17-OCT-2003 (revised)
XX DT 19-DEC-1997 (first entry)
XX XX
XX DE Mouse hyaluronate synthase genomic DNA.
XX KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX OS Mus sp; (strain FM3A).
XX FH Location/Qualifiers
XX FT CDS 49..1800
XX FT /*tag= a
XX FT /product= "Hyaluronate synthase"
XX XX

```

```

DE Mouse hyaluronan synthase (HAS) 1 DNA.
XX KW Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
XX KW eye cell; osteoarthritis; gene therapy; gene; ds.
XX OS Mus sp.
XX FH Location/Qualifiers
XX FT CDS 1..1752
XX FT /*tag= a
XX FT /product= "Mouse hyaluronan synthase 1"
XX PN US2003087850-A1.
XX PD 08-MAY-2003.
XX PF 10-JUL-2001; 2001US-00902939.
XX XX 10-JUL-2001; 2001US-00902939.
XX PA (DEHA/) DEHAZYA P.
XX PA (CHEN/) CHEN W.
XX XX Dehazya P, Chen W;
XX PI WPI; 2003-755151/71.
XX DR P-PSDB; AAE39152.
XX XX Dihydrazone derivatized hyaluronic acid/nucleic acid bioconjugate
PT comprising derivatized hyaluronic acid cross linked to nucleic acid
PT encoding hyaluronan synthase useful for treating dry eye syndrome.
XX PS Claim 1; Page 12-13; 31pp; English.
XX CC The invention relates to dihydrazone derivatized hyaluronic acid (HA)/
CC nucleic acid bioconjugate comprising derivatized HA cross linked to
CC nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
CC for transfecting an eye cell of an individual. It is useful for treating
CC dry eye syndrome and osteoarthritis of the particular joints. The
CC invention is also useful in gene therapy. The present sequence is mouse
CC HAS1 DNA
XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 87.0%; Score 17.4; DB 10; Length 1752;
Best Local Similarity 94.7%; Pred. No. 91;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGCTCTCGAGTAGCA 19
Db 1114 AGGGCGCTCTCGAGTAGCA 1096

RESULT 10
AAT91655/C
ID AAT91655 standard; DNA; 2102 BP.
XX AC AAT91655;
XX AC AAT91655;
XX DT 17-OCT-2003 (revised)
XX DT 19-DEC-1997 (first entry)
XX XX
XX DE Mouse hyaluronate synthase genomic DNA.
XX KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX OS Mus sp; (strain FM3A).
XX FH Location/Qualifiers
XX FT CDS 49..1800
XX FT /*tag= a
XX FT /product= "Hyaluronate synthase"
XX XX

```

PN JP09224674-A.  
 XX  
 PD  
 XX  
 PF 02-SEP-1997.  
 XX  
 PF 26-FEB-1996; 96JP-00038336.  
 XX  
 PF 26-FEB-1996; 96JP-00038336.  
 XX  
 PA (KAGG ) KAGAKU GIJUTSUCHO CHOKAN KANBO.  
 XX  
 DR WPI: 1997-484102/45.  
 DR P-PSDB; AA030704.  
 XX  
 PT Hyaluronate synthase isolated from mouse cells - useful for large-scale  
 PT production of hyaluronic acid.  
 XX  
 PS Claim 2; Page 10-13; 15pp; Japanese.  
 XX  
 CC A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal  
 CC medium containing 10 % heat-inactivated bovine serum, twice concentration  
 CC of amino acids and vitamins and penicillin and streptomycin at 37 degrees  
 CC Celsius. The culture was subjected to immobilised erythrocyte exclusion  
 CC to examine the extent of extracellular formation of hyaluronic acid  
 CC matrix. Cells which showed high formation were recovered and named FM3A  
 CC HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-  
 CC (HAS- P cell ). Poly(A) + RNA was isolated from FM3A HAI and cDNA was  
 CC prepared and was used for constructing a library in HAS- P cells. Cells  
 CC having hyaluronic acid synthetic activity were selected from the  
 CC transfectants and plasmid DNA was recovered and amplified in E.coli. The  
 CC resulting genomic DNA sequence codes for hyaluronate synthase having a  
 CC sequence of 583 amino acids which is used for large-scale production of  
 CC hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;  
 Query Match 87.0%; Score 17.4; DB 2; Length 2102;  
 Best Local Similarity 94.7%; Pred. No. 92;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGGGCGTCTCTGAGTAGCA 19  
 Db 1162 AGGGCGTCTCCGAGTAGCA 1144  
 RESULT 11  
 AAZ10862/c  
 ID AAZ10862 standard; DNA; 2102 BP.  
 AC  
 AC AAZ10862;  
 XX  
 DT 19-OCT-1999 (first entry)  
 DE Hyaluronate synthase coding sequence.  
 XX  
 KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 49..1800  
 FT /\*tag= a  
 FT /product= "HAS1"  
 FT /note= "modified hyaluronate synthase protein"  
 XX  
 PN JP11196875-A.  
 XX  
 PD 27-JUL-1999.  
 XX  
 PF 14-JAN-1998; 98JP-00006191.  
 XX  
 PF 14-JAN-1998; 98JP-00006191.  
 XX  
 PR 14-JAN-1998; 98JP-00006191.  
 XX  
 PA (SEKK ) SEIKAGAKU KOGYO CO LTD.

XX WPI: 1999-496653/42.  
 DR P-PSDB; AAY32503.  
 XX  
 PT New promoter DNA of hyaluronate synthase - used to specifically express  
 PT gene with cell growth inhibiting activity.  
 XX  
 PS Example 1; Page 8-10; 13pp; Japanese.  
 XX  
 CC This sequence represents the mouse hyaluronate synthase coding sequence,  
 CC and was used to isolate the hyaluronate synthase promoter of the  
 CC invention. The promoter can be used for specifically expressing a gene  
 CC having cell growth inhibiting activity  
 XX  
 SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;  
 Query Match 87.0%; Score 17.4; DB 2; Length 2102;  
 Best Local Similarity 94.7%; Pred. No. 92;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGGGCGTCTCTGAGTAGCA 19  
 Db 1162 AGGGCGTCTCCGAGTAGCA 1144  
 RESULT 12  
 AAZ88199/c  
 ID AAZ88199 standard; cDNA to mRNA; 2102 BP.  
 AC  
 AC AAZ88199;  
 XX  
 DT 27-APR-2000 (first entry)  
 DE Mouse hyaluronate synthase modified protein HAS1 encoding cDNA.  
 XX  
 KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;  
 KW research reagent; biochemical research; medical development; ds.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 49..1800  
 FT /\*tag= a  
 FT /product= "HAS1"  
 FT /note= "modified hyaluronate synthase protein"  
 XX  
 PN JP2000004886-A.  
 XX  
 PD 11-JAN-2000.  
 XX  
 PF 24-JUN-1998; 98JP-00193788.  
 XX  
 PF 24-JUN-1998; 98JP-00193788.  
 XX  
 PA (SEKK ) SEIKAGAKU KOGYO CO LTD.  
 XX  
 DR WPI: 2000-140125/13.  
 DR P-PSDB; AAY68491.  
 XX  
 PT A hyaluronate synthase modified protein - useful as a research reagent  
 PT for biochemical research and medical development.  
 XX  
 PS Claim 18; Page 12-14; 30pp; Japanese.  
 XX  
 CC The present invention describes a recombinant protein which consists of  
 CC three continuous regions (N-terminal region, internal region and C-  
 CC terminal region) where one or two regions among the above three regions  
 CC is selected from the three hyaluronate synthase (HAS) modified proteins  
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,  
 CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein  
 CC is useful as a research reagent for biochemical research and medical  
 CC development. The invention provides a HAS modified protein of increased  
 CC or lowered activity. The present sequence encodes mouse HAS1

```
XX SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 87.0%; Score 17.4; DB 3; Length 2102;
Best Local Similarity 94.7%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
DB 1162 AGGGCGTCTCCGAGTAGCA 1144

RESULT 13
AAA39987/c
ID AAA39987 standard; cDNA; 2102 BP.
XX
AC AAA39987;
XX
DT 19-OCT-2000 (first entry)
XX
DE Murine HAS1 cDNA.
XX
KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
KW gene targeting; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /*tag= a
FT /product= "HAS1"
XX
PN JP2000116382-A.
XX
PD 25-APR-2000.
XX
PF 13-OCT-1998; 98JP-00291201.
XX
PR 13-OCT-1998; 98JP-00291201.
XX
PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-369404/32.
DR P-PSDB; AAB09948.
XX
DNA used for gene targeting of the hyaluronate synthase gene (HAS1) .
PS Example; Page 7-9; 14pp; Japanese.
XX
This invention describes a novel DNA which contains an heterologous DNA
(i) to be introduced to chromosomal DNA of a host cell, a first and
second homologous region DNA (II) connected respectively to the 5' and 3'
sides of the introduced DNA, and a negative marker gene (III) expressable
in the host cell. (I) contains a positive marker gene expressable in the
host cell. (I), (II) and the region encoding the intracellular loop of
the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
constituted so as to cause a homologous recombination. The DNA can be
used for gene targeting of the HAS1 gene. This sequence encodes the
murine HAS1 protein described in the method of the invention
XX
Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 87.0%; Score 17.4; DB 3; Length 2102;
Best Local Similarity 94.7%; Pred. No. 92;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19
DB 1162 AGGGCGTCTCCGAGTAGCA 1144

RESULT 14
ACH9039/c
```

```
ID ACH9039 standard; DNA; 239 BP.
XX
AC ACH9039;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #22234.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
WPI; 2004-119264/12.
XX
New human genome-derived single exon nucleic acid probes useful for human
gene expression analysis, for identifying or characterizing alternative
splicing events, for assessing genomic alterations or as tools for
surveying tissues.
XX
Claim 1; SEQ ID NO 22234; 80pp; English.
XX
The invention relates to a nucleic acid probe for measuring human gene
expression, comprising any of the 27,400 fully defined nucleotide
sequences in the specification, or their complements or fragments, and
encoding at least 8 amino acids of any of the 6888 amino acid sequences
fully defined in the specification. The probe is a single exon probe that
hybridises under high stringency conditions to a nucleic acid molecule
expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, and a method of providing
human gene expression data by subscription, and a computer-readable
storage medium which contains a database having a plurality of records
(each record including data on the expression of a single exon probe
cited above. The probe, methods and apparatus are useful in gene
expression analysis. The probes may be used as tools for surveying
tissues to detect the presence of expressed messages that contain their
specific exon or in constructing genome-derived single exon microarrays.
In addition, the probes are used in identifying and characterising
alternative splicing events, in detecting and characterising gross
alterations in the genomic locus that includes their exon, in assessing
smaller genomic alterations, in priming the synthesis of nucleic acids,
or in expressing the ORF-encoded peptide. The present sequence is a human
single exon probe of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
Sequence 239 BP; 54 A; 58 C; 59 G; 68 T; 0 U; 0 Other;
```

```
Query Match      84.0%; Score 16.8; DB 12; Length 239;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTGAGTAGCAG 20
        ||||| ||||| ||||| |||||
Db      113 AGGGCATCTCTGAGTTGCAG 94

RESULT 15
AAC10010
ID AAC10010 standard; cDNA; 278 BP.
XX
AC AAC10010;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 14085.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 14085; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 278 BP; 74 A; 70 C; 65 G; 69 T; 0 U; 0 Other;

Query Match      84.0%; Score 16.8; DB 3; Length 278;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTGAGTAGCAG 20
        ||||| ||||| ||||| |||||
Db      113 AGGGCATCTCTGAGTTGCAG 132

RESULT 16
AAC02292
ID AAC02292 standard; cDNA; 313 BP.
XX
AC AAC02292;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 2290.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PD 06-SEP-2000.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 2290; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX
SQ Sequence 313 BP; 83 A; 78 C; 72 G; 76 T; 0 U; 4 Other;

Query Match      84.0%; Score 16.8; DB 3; Length 313;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTGAGTAGCAG 20
        ||||| ||||| ||||| |||||
Db      113 AGGGCATCTCTGAGTTGCAG 132

RESULT 17
ACH75339/c
ID ACH75339 standard; DNA; 595 BP.
XX
AC ACH75339;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #8534.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
```

```
OS Homo sapiens.
XX US2003194704-A1.
XX 16-OCT-2003.
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX New human genome-derived single exon nucleic acid probes useful for human
XX gene expression analysis, for identifying or characterizing alternative
XX splicing events, for assessing genomic alterations or as tools for
XX surveying tissues.
XX Claim 15; SEQ ID NO 8534; 80pp; English.
XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
XX sequences in the specification, or their complements or fragments, and
XX encoding at least 8 amino acids of any of the 6888 amino acid sequences
XX fully defined in the specification. The probe is a single exon probe that
XX hybridises under high stringency conditions to a nucleic acid molecule
XX expressed in human cells or tissues. Also included are a spatially-
XX addressable set of single exon nucleic acid probes for measuring human
XX gene expression (comprising a plurality of single exon nucleic acid
XX probes cited above, where each of the plurality of probes is separately
XX and addressably isolatable or amplifiable from the plurality), a single
XX exon microarray for measuring human gene expression, a method of
XX measuring human gene expression, a vector comprising the single exon
XX probe cited above, an ORF-encoded peptide comprising at least 8
XX contiguous amino acids of any of the above-mentioned amino acid
XX sequences (optionally with conservative amino acid substitutions), an
XX isolated antibody that binds specifically to a peptide cited above,
XX a customer desiring to measure gene expression, a method of providing
XX human gene expression data by subexpression, and a computer-readable
XX storage medium which contains a database having a plurality of records
XX (each record including data on the expression of a single exon probe
XX cited above. The probe, methods and apparatus are useful in gene
XX expression analysis. The probes may be used as tools for surveying
XX tissues to detect the presence of expressed messages that contain their
XX specific exon, or in constructing genome-derived single exon microarrays.
XX In addition, the probes are used in identifying and characterising
XX alternative splicing events, in detecting and characterising gross
XX alterations in the genomic locus that includes their exon, in assessing
XX smaller genomic alterations, in priming the synthesis of nucleic acids
XX or in expressing the ORF-encoded peptide. The present sequence is a human
XX single exon probe of the invention. Note: The present sequence for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030194704
XX Sequence 595 BP; 145 A; 163 C; 127 G; 160 T; 0 U; 0 Other;
XX Query Match 84.0%; Score 16.8; DB 12; Length 595;
XX Best Local Similarity 90.0%; Pred. No. 1.6e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 422 AGGGCATCTCTGAGTTGCAG 403
XX
XX RESULT 18
XX Query Match 84.0%; Score 16.8; DB 12; Length 595;
XX Best Local Similarity 90.0%; Pred. No. 1.6e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 422 AGGGCATCTCTGAGTTGCAG 403
XX
XX RESULT 19
XX ABID33213
XX ID ABD33213 standard; DNA; 80105 BP.
XX AC ABD33213;
XX AC ABD33213;
XX DT 18-NOV-2004 (first entry)
XX DE Murine cancer-associated (CA) gene MD07-034.
XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX ds; cancer; cytostatic.
XX OS Mus musculus.
XX PN WO2004058146-A2.
XX PD 15-JUL-2004.
XX PF 15-DEC-2003; 2003WO-US040081.
XX PR 17-DEC-2002; 2002US-00322281.
```

```
AAH19499
ID AAH19499 standard; cDNA; 775 BP.
XX AC AAH19499;
XX DT 03-AUG-2001 (first entry)
XX DE Human galectin 15 coding sequence.
XX KW Human; galectin 15; cytostatic; antiinflammatory; immunomodulatory;
XX haemostatic; anti-HIV; gene therapy; malignant tumour; haemopathy;
XX HIV infection; immunological disease; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200130837-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-CN000375.
XX PR 28-OCT-1999; 99CN-00119900.
XX (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.
XX Mao Y, Xie Y;
XX WPI; 2001-291028/30.
XX P-PSDB; AAB84886.
XX Human galectin 15 and its encoding polynucleotides are applicable in
XX diagnosis and treatment of malignant tumors, hemopathy, HIV infection,
XX immunological diseases and various inflammations.
XX Claim 6; Page 21; 28pp; Chinese.
XX The present sequence is the coding sequence for human galectin 15.
XX Galectin 15 is useful in diagnosis and treatment of malignant tumour,
XX haemopathy, HIV infection, immunological diseases and inflammation
XX Sequence 775 BP; 209 A; 195 C; 170 G; 201 T; 0 U; 0 Other;
XX Query Match 84.0%; Score 16.8; DB 5; Length 775;
XX Best Local Similarity 90.0%; Pred. No. 1.7e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 128 AGGGCATCTCTGAGTTGCAG 147
XX
XX RESULT 19
XX ABID33213
XX ID ABD33213 standard; DNA; 80105 BP.
XX AC ABD33213;
XX AC ABD33213;
XX DT 18-NOV-2004 (first entry)
XX DE Murine cancer-associated (CA) gene MD07-034.
XX KW Mouse; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
XX ds; cancer; cytostatic.
XX OS Mus musculus.
XX PN WO2004058146-A2.
XX PD 15-JUL-2004.
XX PF 15-DEC-2003; 2003WO-US040081.
XX PR 17-DEC-2002; 2002US-00322281.
```

XX (SAGR-) SAGRES DISCOVERY INC.  
 XX Morris DW, Malandro MS;  
 XX WPI; 2004-499109/47.  
 XX Novel human cancer associated protein encoded within open reading frame  
 XX of cancer associated gene, useful as targets for diagnosing cancer.  
 XX Disclosure; SEQ ID NO 223; 182pp; English.  
 XX The invention relates to cancer-associated proteins (CAP) and the cancer-  
 XX associated (CA) nucleic acids encoding them. The invention also relates  
 XX to a method for treating cancers involving administering to a patient an  
 XX inhibitor of CAP, and a method of screening for anticancer activity in a  
 XX potential drug involving providing a cell that expresses a CA gene,  
 XX contacting a tissue sample derived from a cancer cell with an anticancer  
 XX drug candidate and monitoring the effect of the anticancer drug candidate  
 XX on expression of the CA gene. The CAP proteins are useful for detecting  
 XX cancer associated with expression of a CAP protein in a test cell sample  
 XX and for screening for a bioactive agent capable of modulating the  
 XX activity of a CAP protein. The CA nucleic acids are useful for diagnosing  
 XX cancer, involving determining the expression of a CA nucleic acid in a  
 XX tissue. This sequence represents a murine CA gene of the invention. Note:  
 XX The sequence data for this patent did not form part of the printed  
 XX specification, but was obtained in electronic format directly from WIPO  
 XX at ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 80105 BP; 17052 A; 21395 C; 22145 G; 18961 T; 0 U; 552 Other;  
 XX Query Match 84.0%; Score 16.8; DB 13; Length 80105;  
 XX Best Local Similarity 90.4%; Pred. No. 2.7e+02;  
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AGGGCTCTCTGAGTAGCAG 20  
 DB 32700 AGGGCTCTCTGAGTAGCAG 32719  
 RESULT 20  
 AX98492  
 ID AAX98492 standard; cDNA; 300 BP.  
 XX AAX98492;  
 XX 24-SEP-1999 (first entry)  
 XX Human cancer cell derived cDNA #218.  
 XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;  
 XX integral membrane protein; aspartyl protease; GATA family; wt family;  
 XX transcription factor; G-protein alpha subunit; protein phosphatase;  
 XX phospholipase C; tyrosine phosphatase; diacylglycerol binding protein; trypsin;  
 XX protein kinase; tyrosine phosphatase; developmental signalling protein;  
 XX WW/rps5/WMP domain; therapy; forensic; genetic mapping; diagnostic;  
 XX detection; treatment; cervical; melanoma; colorectal adenocarcinoma;  
 XX Wilm's tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;  
 XX leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;  
 XX prostate; ss.  
 XX Homo sapiens.  
 XX WO9333982-A2.  
 XX 08-JUL-1999.  
 XX 22-DEC-1998; 98WO-US027610.  
 XX 23-DEC-1997; 97US-0068755P.  
 XX 03-APR-1998; 98US-0080664P.  
 XX 21-OCT-1998; 98US-0105234P.  
 XX 27-OCT-1998; 98US-0105877P.

PR 21-DEC-1998; 98US-00217471.  
 XX (CHIR) CHIRON CORP.  
 XX (HYSE-) HYSEQ INC.  
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
 XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
 XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
 XX Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;  
 XX WPI; 1999-430243/36.  
 XX New isolated human polynucleotides.  
 XX Claim 1; Page 370; 591pp; English.  
 XX This invention describes novel isolated human polynucleotides obtained by  
 XX screening for differential expression in colon cancer, breast cancer and  
 XX lung cancer cell lines. The polynucleotides of the invention are  
 XX represented in AAX98275-X99118 and encode polypeptides of protein  
 XX families selected from 4 transmembrane segments integral membrane  
 XX proteins, 7 transmembrane receptors, ATPases associated with various  
 XX cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of  
 XX transcription factors, G-protein alpha subunit, phospholipase C or  
 XX diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,  
 XX protein tyrosine phosphatase, trypsin, wt family of developmental  
 XX signalling proteins and WW/rps5/WMP domain containing proteins. The  
 XX encoded polypeptides also have a functional domain selected from Ank  
 XX repeat, basic region plus leucine zipper transcription factors,  
 XX bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger  
 XX (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease  
 XX domain. The polynucleotides encode polypeptides with similarity to known  
 XX protein families and are predicted to have similar properties. The novel  
 XX polynucleotides can be used to develop products for use as therapeutic  
 XX agents and in forensics, genetic analysis, mapping and diagnostic  
 XX applications. In particular, the product can be used for the detection  
 XX and management of cancers. They can be used for treating e.g. cervical  
 XX cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas,  
 XX retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic  
 XX myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and  
 XX hereditary ectodermal dysplasia, congenital alveolar dysplasia,  
 XX epithelial dysplasia of the cervix, fibrous dysplasia of bone, and  
 XX mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,  
 XX prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of  
 XX the skin  
 XX SQ Sequence 300 BP; 58 A; 88 C; 99 G; 55 T; 0 U; 0 Other;  
 Query Match 82.0%; Score 16.4; DB 2; Length 300;  
 Best Local Similarity 94.4%; Pred. No. 2.4e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GGCGTCTCTGAGTAGCAG 20  
 DB 222 GGCGTCTCTGAGTAGCAG 239  
 RESULT 21  
 AAX99116  
 ID AAX99116 standard; cDNA; 452 BP.  
 XX AAX99116;  
 XX 24-SEP-1999 (first entry)  
 XX Human cancer cell derived cDNA contig #42.  
 XX Cancer; human; colon; breast; lung; transmembrane receptor; ATPase;  
 XX integral membrane protein; aspartyl protease; GATA family; wt family;  
 XX transcription factor; G-protein alpha subunit; protein phosphatase;  
 XX phospholipase C; tyrosine phosphatase; diacylglycerol binding protein; trypsin;  
 XX protein kinase; tyrosine phosphatase; developmental signalling protein;

KW WW/rps5/WWP domain; therapy; forensic; genetic mapping; diagnostic;  
KW detection; treatment; cervical; melanoma; colorectal adenocarcinoma;  
KW Wilms' tumour; retinoblastoma; sarcoma; myosarcoma; lung carcinoma;  
KW leukemia; lymphoma; dysplasia; hyperplasia; endometrium; adrenal;  
KW prostate; ss.  
XX Homo sapiens.  
XX  
XX  
XX WO9933982-A2.  
XX  
XX  
XX 08-JUL-1999.  
XX  
XX 22-DEC-1998; 98WO-US027610.  
XX  
XX 23-DEC-1997; 97US-0068755P.  
XX 03-APR-1998; 98US-0080664P.  
XX 21-OCT-1998; 98US-0105234P.  
XX 27-OCT-1998; 98US-0105977P.  
XX 21-DEC-1998; 98US-00217471.  
XX  
XX (CHIR ) CHIRON CORP.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kaseam A;  
PI Lamson G, Drmanac R, Ckvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;  
XX WPI; 1999-430243/36.  
XX  
XX  
XX New isolated human polynucleotides.  
XX  
XX  
XX Claim 1; Page 590; 591pp; English.  
XX  
XX This invention describes novel isolated human polynucleotides obtained by  
XX screening for differential expression in colon cancer, breast cancer and  
XX lung cancer cell lines. The polynucleotides of the invention are  
XX represented in AA98275-X99118 and encode polypeptides of protein  
XX families selected from 4 transmembrane segments integral membrane  
XX proteins, 7 transmembrane receptors, Arpases associated with various  
XX cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of  
XX transcription factors, G-protein alpha subunit, phorbolsters or  
XX diacylglycerol binding proteins, protein kinase, protein phosphatase 2C,  
XX protein tyrosine phosphatase, trypsin, wnt family of developmental  
XX signalling proteins and WW/rps5/WWP domain containing proteins. The  
XX encoded polypeptides also have a functional domain selected from Ank  
XX repeat, basic region plus leucine zipper transcription factors,  
XX bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger  
XX (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease  
XX domain. The polynucleotides encode polypeptides with similarity to known  
XX protein families and are predicted to have similar properties. The novel  
XX polynucleotides can be used to develop products for use as therapeutic  
XX agents and in forensics, genetic analysis, mapping and diagnostic  
XX applications. In particular, the product can be used for the detection  
XX and management of cancers. They can be used for treating e.g. cervical  
XX cancers, melanomas, colorectal adenocarcinomas, Wilms' tumour, sarcomas,  
XX retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic  
XX myelogenous leukemia, promyelocytic leukemia, monocytic leukemia, and  
XX myeloid leukemia, and lymphomas such as histiocytic lymphoma, anhydric  
XX hereditary ectodermal dysplasia, congenital alveolar dysplasia,  
XX epithelial dysplasia of the cervix, fibrous dysplasia of bone, and  
XX mammary dysplasia, hyperplasias, e.g. endometrial, adrenal, breast,  
XX prostate or thyroid hyperplasias or pseudoepitheliomatous hyperplasia of  
XX the skin  
XX  
XX Sequence 452 BP; 102 A; 120 C; 146 G; 79 T; 0 U; 5 Other;  
XX  
XX  
XX Query Match 82.0%; Score 16.4; DB 2; Length 452;  
XX Best Local Similarity 94.4%; Pred. No. 2.5e+02;  
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 3 GGCGTCTCTGAGTAGCAG 20  
XX |||||||||||||||||

DB 222 GGCGTCTCTGAGTAGCAG 239  
RESULT 22  
ID ADJ42940/c  
XX ADJ42940 standard; cDNA; 748 BP.  
XX  
XX ADJ42940;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Plant cDNA #3940.  
XX  
XX Plant; gene; ss; transcription; plant genome augmentation; cereal;  
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;  
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;  
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;  
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;  
KW antifungal.  
XX  
XX Eukaryota.  
XX  
XX US2004016025-A1.  
XX  
XX 22-JAN-2004.  
XX  
XX 26-SEP-2002; 2002US-00260239.  
XX  
XX 26-SEP-2001; 2001US-0325277P.  
XX 26-SEP-2001; 2001US-0325448P.  
XX 04-APR-2002; 2002US-0370620P.  
XX  
XX (BUDW/) BUDWORTH P.  
XX (MOUG/) MOUGHAMER T.  
XX (BRIG/) BRIGGS S P.  
XX (COOP/) COOPER B.  
XX (GLAZ/) GLAZEBROOK J.  
XX (GOF/) GOFF S A.  
XX (KATA/) KATAGIRI F.  
XX (KREP/) KREPS J.  
XX (PROV/) PROVART N.  
XX (RICK/) RICHE D.  
XX (ZHUT/) ZHU T.  
XX  
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;  
XX Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;  
XX WPI; 2004-190374/18.  
XX  
XX New rice promoter, useful for manipulating crop plants to alter or  
XX improve phenotypic characteristics, e.g. produce large quantities of oil  
XX or proteins, resistance to insecticides, virus or fungi, stress tolerance  
XX or high nutritional value.  
XX  
XX Example 13; SEQ ID NO 3940; 230pp; English.  
XX  
XX The invention relates to plant nucleotide sequences that direct seed-,  
XX leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential  
XX or constitutive transcription of an operatively linked nucleic acid  
XX segment. The invention also relates to a method for augmenting a plant  
XX genome and a method of identifying a gene, where its expression is  
XX altered in the seed, leaf, stem, panicle, pollen, root or is constitutive  
XX in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,  
XX canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,  
XX sorghum, rice or wheat. The polynucleotides and the polypeptides they  
XX encode are useful for manipulating crop plants to alter or improve  
XX phenotypic characteristics, to produce large quantities of oil or  
XX proteins, to incur resistance to insecticides, viruses or fungi, and to  
XX incur stress tolerance (e.g. salt, cold or drought) to ensure the plants  
XX have a high nutritional value with reduced apical dominance or dwarfism,  
XX early flowering or altered metabolic pathways. This sequence represents a  
XX plant nucleic acid of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification but was obtained in



CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 748 BP; 170 A; 209 C; 194 G; 175 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 748;  
Best Local Similarity 94.4%; Pred. No. 2.6e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGTCTCTCGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 235 GCGTCTCTCGAGTAGCAG 218

RESULT 23  
ADB62940  
ID ADB62940 standard; cDNA; 1611 BP.  
XX  
AC ADB62940;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human cDNA encoding clone PLACB60132320.  
XX  
KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;  
KW tissue regeneration; cell regeneration; membrane protein;  
KW signal transduction-related protein; transcription-related protein;  
KW osteoporosis; neurological disease; cancer; tumour.  
XX  
OS Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT CDS 471..1121  
FT /tag= a  
FT /product= "Clone PLACB60132320 protein"  
XX

PN EP1308459-A2.  
XX  
PD 07-MAY-2003.  
XX  
PF 28-MAR-2002; 2002EP-00007401.  
XX  
PR 05-NOV-2001; 2001JP-00379298.  
PR 25-JAN-2002; 2002US-00350978.  
XX  
PA (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-450961/43.  
DR P-PSDB; ADB64910.

XX New polynucleotides and polypeptides, useful for developing a diagnostic  
PT marker or medicines for regulation of their expression and activity, or  
PT as targets of gene therapy.  
XX  
PS Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected  
CC from 1970 fully defined nucleotide sequences which encode novel  
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
CC or its partial peptide, an antibody binding to the polypeptide or peptide  
CC of the polynucleotide, immunologically assaying the polypeptide or  
CC peptide of the polynucleotide by contacting the polypeptide or peptide  
CC with the antibody of the encoded protein, and observing the binding  
CC between the two, a transformant carrying the polynucleotide in an  
CC expressible manner and an antisense polynucleotide. The oligonucleotide  
CC is useful as a primer for synthesising the polynucleotide, or as a probe  
CC for detecting the polynucleotide. The polynucleotides and encoded  
CC proteins are useful as pharmaceutical agents and many disease-related  
CC genes may be included in them, for developing a diagnostic marker or

CC medicines for regulation of their expression and activity, or as targets  
CC of gene therapy. The genes are involved in tissue and/or cell  
CC regeneration. Membrane proteins, signal transduction-related proteins,  
CC transcription-related proteins, disease-related proteins and genes  
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate  
CC the activity or expression of the encoded protein to treat diseases. The  
CC sequence presented is a cDNA of the invention. Note: Some of the sequence  
CC data for this patent is not represented in the printed specification, but  
CC is based on sequence information supplied by the European Patent Office.  
XX  
SQ Sequence 1611 BP; 333 A; 435 C; 516 G; 327 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 1611;  
Best Local Similarity 94.4%; Pred. No. 2.9e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 636 GCGTCTCTGAGTAGCAG 653

RESULT 24  
AAC77862  
ID AAC77862 standard; cDNA; 1992 BP.  
XX  
AC AAC77862;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human cancer associated gene sequence SEQ ID NO:256.

XX  
KW Human; cancer associated gene; cancer antigen; detection; cancer;  
KW diagnosis; cytostatic; proliferative; cancer; immunomodulator;  
KW antidiabetic; antiaesthetic; antineumatic; antineumatic; antiviral;  
KW antineumatic; antineumatic; antineumatic; antineumatic; antineumatic;  
KW dermatological; neuroprotective; thrombolytic; coagulant; neurologic;  
KW vasotropic; antipsoriatic; angiogenic; gene therapy; inflammation;  
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
KW allergic reaction; graft versus host disease; organ rejection;  
KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
KW neurological disease; drug screening; ss.

XX Homo sapiens.  
XX WO200055350-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 08-MAR-2000; 2000WO-US005882.  
XX  
PR 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587533/55.  
DR P-PSDB; AAB43653.

XX Novel isolated nucleic acids comprising sequences encoding peptides  
PT useful for treating or diagnosing e.g. cancer.  
XX  
PS Claim 1; Page 821-822; 2352pp; English.  
XX  
CC AAC77607 to AAC78448 encode the human cancer associated proteins given in  
CC AAB43398 to AAB44239. The proteins can have activities based on the  
CC tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; cancer; immunomodulator;  
CC antidiabetic; antiaesthetic; antineumatic; antineumatic; antineumatic;  
CC antineumatic; antineumatic; antineumatic; antineumatic; antineumatic;  
CC dermatological; neuroprotective; thrombolytic; coagulant; neurologic;  
CC vasotropic; antipsoriatic and angiogenic. The

CC	polynucleotides and polypeptides can be used for preventing, treating or	CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	ameliorating medical conditions and diagnosing pathological conditions.	CC	polynucleotide which comprises a 3'-end sequence, where the
CC	Polynucleotides, polypeptides, antibodies, agonists and antagonists from	CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the present invention may be used to treat immune disorders by activating	CC	the 5'-end sequence/3'-end sequence is selected from those defined in the
CC	or inhibiting the proliferation, differentiation or mobilisation of	CC	specification. The primer sets can be used in antisense therapy and in
CC	immune cells, to treat disorders of haematopoietic cells, autoimmune	CC	gene therapy. The primers are useful for synthesising polynucleotides,
CC	disorders, allergic reactions, graft versus host disease and organ	CC	particularly full-length cDNAs. The primers are also useful for the
CC	rejection, modulate haemostatic or thrombolytic activity, modulate	CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	inflammation, cancers, cardiovascular disorders, neurological disease and	CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	bacterial or viral infections. The peptides, nucleotides, antibodies,	CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	agonists and antagonists may be also be used in drug screens. AAC78449 to	CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	AAC78457 and AAB44240 represent sequences used in the exemplification of	CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	the present invention	CC	oligonucleotides, all of which are used in the exemplification of the
XX		CC	present invention
SQ	Sequence 1992 BP; 396 A; 573 C; 628 G; 389 T; 0 U; 6 Other;	SQ	Sequence 2057 BP; 409 A; 594 C; 642 G; 412 T; 0 U; 0 Other;
Query Match 82.0%; Score 16.4; DB 3; Length 1992;			
Best Local Similarity 94.4%; Pred. No. 2.9e+02;			
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	3 GCGCTCTCTGAGTAGCAG 20	QY	3 GCGCTCTCTGAGTAGCAG 20
Db	631 GCGCTCTCTGAGCAGCAG 648	Db	1082 GCGCTCTCTGAGCAGCAG 1099
RESULT 25			
AAH17632		ADP83544	
ID	AAH17632 standard; cDNA; 2057 BP.	ID	ADP83544 standard; cDNA; 2250 BP.
XX		XX	
AC	AAH17632;	AC	ADP83544;
XX		XX	
DT	26-JUN-2001 (first entry)	DT	26-FEB-2004 (first entry)
XX		DE	Human GRIM1 cDNA.
DE	Human cDNA sequence SEQ ID NO:17166.	XX	GRIM1; global repressor involved in myogenic differentiation; vulnery;
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.	KW	muscular degeneration; muscle formation; skeletal muscle differentiation;
XX		KW	fat cell differentiation; degenerative muscle disease; muscle injury; ss;
OS	Homo sapiens.	XX	gene; chromosome lp33.36.
PN	EP1074617-A2.	XX	Homo sapiens.
XX		XX	
PD	07-FEB-2001.	XX	
XX		XX	
PF	28-JUL-2000; 2000EP-00116126.	XX	
XX		XX	
PR	29-JUL-1999; 99JP-00248036.	XX	
PR	27-AUG-1999; 99JP-00300253.	XX	
PR	11-JAN-2000; 2000JP-00118776.	XX	
PR	02-MAY-2000; 2000JP-00183767.	XX	
PR	09-JUN-2000; 2000JP-00241899.	XX	
XX		XX	
PA	(HELI-) HELIX RES INST.	XX	
XX		XX	
PI	Ota T, Isogai T, Nishikawa T, Hayaehi K, Saito K, Yamamoto J;	XX	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	XX	
XX		XX	
DR	WPI; 2001-318749/34.	XX	
XX		XX	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-	XX	
PT	length cDNAs defined in the specification, and for the detection and/or	XX	
PT	diagnosis of the abnormality of the proteins encoded by the full-length	XX	
PT	cDNAs.	XX	
XX		XX	
PS	Claim 8; SEQ ID NO 17166; 2537pp + Sequence Listing; English.	XX	
XX		XX	
CC	The present invention describes primer sets for synthesising 5602 full-	XX	
CC	length cDNAs defined in the specification. Where a primer set comprises:	XX	
CC	(a) an oligo-dr primer and an oligonucleotide complementary to the	XX	
CC	complementary strand of a polynucleotide which comprises one of the 5602	XX	
CC	nucleotide sequences defined in the specification, where the	XX	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	XX	
CC	of an oligonucleotide comprising a sequence complementary to the	XX	
CC	complementary strand of a polynucleotide which comprises a 5'-end	XX	

CC murine sequence has been identified by comparative screening of murine  
 CC expressed sequence tags (EST's). The products of the invention have  
 CC vulnary activity and the method of the invention can be used to modify  
 CC GRIM function, by blocking movement of GRIM between nucleus and  
 CC cytoplasm, to counteract existing muscular degeneration or activate  
 CC muscle formation. GRIM is a repressor of many complexes and promoters.  
 CC The GRIM polypeptides are used to treat disorders of skeletal muscle  
 CC differentiation or fat cell differentiation, including degenerative  
 CC muscle diseases and muscle injuries in the elderly and to identify  
 CC specific modulators. This sequence encodes the human GRIM1 protein.  
 XX  
 XX Sequence 2250 BP; 490 A; 639 C; 720 G; 401 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 2250;  
 Best Local Similarity 94.4%; Pred. No. 3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 1780 GCGCTCTCTGAGTAGCAG 1797

## RESULT 27

ACN37249

ID ACN37249 standard; cDNA; 2508 BP.

XX ACN37249;

XX 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) cDNA DNA323721, SEQ ID NO:6.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
 KW central nervous system cancer; bladder cancer; pancreatic cancer;  
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
 KW chromosome identification; chromosome mapping; gene mapping;  
 KW gene therapy; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

XX 02-OCT-2002; 2002US-0414971P.

XX (GETH ) GENENTECH INC.

XX Wu TD, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX P-PSDB; ABM80001.

XX New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.

XX Claim 1; SEQ ID NO 6; 7273bp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a

CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with  
 CC increased TAT expression, particularly cancers such as breast cancer,  
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
 CC used as hybridisation probes, in chromosome and gene mapping, in  
 CC chromosome identification and in gene therapy. The present sequence  
 CC represents a TAT nucleic acid of the invention

XX Sequence 2508 BP; 534 A; 705 C; 782 G; 487 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 13; Length 2508;

Best Local Similarity 94.4%; Pred. No. 3e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20

|||||

Db 1548 GCGCTCTCTGAGTAGCAG 1565

## RESULT 28

ABL90233/c

ID ABL90233 standard; cDNA; 2665 BP.

XX ABL90233;

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 795.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US016450.

XX 19-MAY-2000; 2000US-0205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

XX P-PSDB; ABB89824.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 PT disorders.

XX Claim 4; SEQ ID NO 795; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)  
 CC  
 XX  
 SQ Sequence 2665 BP; 518 A; 846 C; 746 G; 550 T; 0 U; 5 Other;  
 Query Match 82.0%; Score 16.4; DB 6; Length 2665;  
 Best Local Similarity 94.4%; Pred. No. 3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 1169 GCGCTCTCTGAGTAGCAG 1152  
 RESULT 29  
 AAH18559  
 ID AAH18559 standard; cDNA; 2753 BP.  
 XX  
 AC AAH18559;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:18728.  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-00116126.  
 PR  
 XX 29-JUL-1999; 99JP-00248036.  
 PR 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX  
 PS Claim 8; SEQ ID NO 18728; 2537pp + Sequence Listing; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dt primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides,  
 CC

CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences. AAH92446 to AAH95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 XX  
 SQ Sequence 2753 BP; 580 A; 775 C; 872 G; 526 T; 0 U; 0 Other;  
 Query Match 82.0%; Score 16.4; DB 4; Length 2753;  
 Best Local Similarity 94.4%; Pred. No. 3e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 1793 GCGCTCTCTGAGTAGCAG 1810  
 RESULT 30  
 AAH72382  
 ID AAH72382 standard; cDNA; 2781 BP.  
 XX  
 AC AAH72382;  
 XX  
 DT 19-DEC-2000 (first entry)  
 XX  
 DE Human nucleic acid-binding protein NuABP-1 cDNA.  
 KW Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;  
 KW expressed sequence tag; drug screening; recombinant expression; antibody;  
 KW reproductive disorder; infertility; immunological disorder;  
 KW neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200004900-A2.  
 PD  
 XX 03-AUG-2000.  
 XX  
 PF 28-JAN-2000; 2000WO-US002237.  
 XX  
 PR 29-JAN-1999; 99US-0117904P.  
 PR 29-JAN-1999; 99US-0117905P.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Tang YT, Lal P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;  
 XX  
 DR WPI; 2000-499332/44.  
 DR P-PSDB; AAB20997.  
 XX  
 PT Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer.  
 XX  
 PS Claim 4; Page 145; 180pp; English.  
 XX  
 CC Sequences AAH72382-A72436 represent cDNAs encoding novel human nucleic  
 CC acid-binding proteins (NuABPs; AAB20997-B21051). These cDNAs were  
 CC produced by extension from an appropriate EST (expressed sequence tag)  
 CC using primers designed using the EST. The invention also relates to  
 CC expression constructs, host cells and transgenic organisms comprising a  
 CC human NuABP nucleic acid, recombinant production of the human NuABPs, and  
 CC antibodies against the human NuABPs, and also to methods of screening  
 CC modulators of human NuABP activity or expression. The human NuABPs, and  
 CC their agonists and antagonists are used to treat diseases associated with  
 CC overexpression or underexpression of functional NuABPs. Human NuABP  
 CC proteins and nucleotides, and NuABP agonists and antagonists can be used  
 CC to diagnose, treat and prevent reproductive, immunological, neurological  
 CC and cell proliferative disorders. Reproductive disorders that may be

CC treated using compositions of the invention include infertility,  
 CC endometriosis, disruptions of the menstrual cycle and disruptions of  
 CC spermatogenesis. Immunological disorders that may be treated include  
 CC AIDS, allergies, and autoimmune disorders such as multiple sclerosis,  
 CC rheumatoid arthritis, diabetes and systemic lupus erythematosus.  
 CC Neurological disorders that may be treated include epilepsy, and  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, prion diseases such as Creutzfeldt-Jakob disease, and mental  
 CC disorders such as schizophrenia. Cell proliferative disorders that may be  
 CC treated include a wide variety of cancers, and also arteriosclerosis,  
 CC atherosclerosis, cirrhosis and psoriasis  
 XX  
 SQ Sequence 2781 BP; 594 A; 779 C; 881 G; 527 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 3; Length 2781;  
 Best Local Similarity 94.4%; Pred. No. 3.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 1810 GCGCTCTCTGAGTAGCAG 1827

RESULT 31  
 AAK52536/c  
 ID AAK52536 standard; cDNA; 3163 BP.

XX AC AAK52536;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 2065.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.  
 XX  
 PN WO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US004098.  
 XX  
 PR 03-FEB-2000; 2000US-00496914.  
 PR 27-APR-2000; 2000US-00560875.  
 PR 20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00620325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00663561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI; 2001-476283/51.  
 DR P-PSDB; AAM79403.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.

XX  
 PS Claim 1; Page 4450-4451; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 SQ Sequence 3163 BP; 611 A; 1008 C; 900 G; 644 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 4; Length 3163;  
 Best Local Similarity 94.4%; Pred. No. 3.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 1349 GCGCTCTCTGAGTAGCAG 1332

RESULT 32  
 AAK51552  
 ID AAK51552 standard; cDNA; 3236 BP.

XX AC AAK51552;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 97.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.  
 XX  
 PN WO200157190-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US004098.  
 XX  
 PR 03-FEB-2000; 2000US-00496914.  
 PR 27-APR-2000; 2000US-00560875.  
 PR 20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00620325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00663561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI; 2001-476283/51.  
 DR P-PSDB; AAM78419.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.

XX  
 PS Claim 1; Page 717-720; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 SQ Sequence 3236 BP; 656 A; 924 C; 1040 G; 613 T; 0 U; 3 Other;  
 Query Match 82.0%; Score 16.4; DB 4; Length 3236;  
 Best Local Similarity 94.4%; Pred. No. 3.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 DB 1872 GCGCTCTCTGAGTAGCAG 1889  
 RESULT 33  
 AAC75812  
 ID AAC75812 standard; cDNA; 3619 BP.  
 XX  
 AC AAC75812;  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1367 polynucleotide sequence SEQ ID NO:2733.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipapillary; antipapillary; antipapillary; antipapillary;  
 KW anticonvulsant; osteopathic; antipapillary; antipapillary; antipapillary;  
 KW immunosuppressant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antidiabetic; antidiabetic;  
 KW thrombosis; contraceptive; SS.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US008621.  
 XX  
 PR 31-MAR-1999; 99US-0127607P.  
 PR 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Shimketa RA, Leach M;  
 PI  
 XX WPI; 2000-602362/57.  
 DR  
 DR P-PSDB; AAB41603.  
 XX  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 XX Claim 5; Page 1974-1976; 5507pp; English.  
 PS  
 XX AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;

CC antipapillary; antipapillary; antipapillary; antipapillary;  
 CC anticonvulsant; antipapillary; antipapillary; antipapillary;  
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antidiabetic; antidiabetic;  
 CC antiviral; antifungal; antirheumatic; antithyroid; antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antidiabetic disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 3619 BP; 718 A; 1049 C; 1153 G; 696 T; 0 U; 3 Other;  
 Query Match 82.0%; Score 16.4; DB 3; Length 3619;  
 Best Local Similarity 94.4%; Pred. No. 3.1e+02;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 GCGCTCTCTGAGTAGCAG 20  
 |||||  
 DB 2029 GCGCTCTCTGAGTAGCAG 2046  
 RESULT 34  
 ADQ23383/c  
 ID ADQ23383 standard; DNA; 4560 BP.  
 XX  
 AC ADQ23383;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 6203.  
 XX  
 KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
 KW ds.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2004048938-A2.  
 PN  
 XX 10-JUN-2004.  
 PD  
 PF 26-NOV-2003; 2003WO-US038193.  
 XX  
 PF 26-NOV-2002; 2002US-0429739P.  
 XX  
 PR (PROT-) PROTEIN DESIGN LABS INC.  
 XX  
 PI Aziz N, Ginsburg WM, Zlotnik A;  
 XX  
 XX WPI; 2004-441208/41.  
 DR  
 XX  
 XX Early detection of soft tissue sarcoma comprises determining expression  
 PT of a gene in a first soft tissue sample and a normal soft tissue sample  
 PT and comparing the gene expression, also useful in treating soft tissue  
 PT sarcoma.  
 XX  
 XX Example 2; SEQ ID NO 6203; 210pp; English.  
 PS  
 XX The invention relates to a novel method for detecting soft tissue sarcoma  
 CC which comprises obtaining a first soft tissue sample from an individual  
 CC and a normal soft tissue sample from the same or different individual,  
 CC determining the expression of a gene in both samples and comparing the  
 CC expression of the gene in both soft tissue samples, where a higher level  
 CC of protein expression in the first soft tissue sample indicates the  
 CC presence of soft tissue sarcoma. The method of the invention has  
 CC cytostatic applications and may be useful for detecting soft tissue

CC sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. CC The current sequence is that of a human soft tissue sarcoma-upregulated CC DNA of the invention. The current sequence is not shown within the CC specification per se but was submitted in CD format by the inventor. XX  
SQ Sequence 4560 BP; 803 A; 1537 C; 1370 G; 850 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 4560;  
Best Local Similarity 94.4%; Pred. No. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20  
DB 2727 GCGCTCTCTGAGTAGCAG 2710  
|||||

RESULT 35  
ADP83532  
ID ADF83532 standard; DNA; 15105 BP.  
XX  
AC ADF83532;  
XX  
DT 26-FEB-2004 (first entry)  
XX  
DE Human GRIM1 DNA.  
XX  
KW GRIM-1; global repressor involved in myogenic differentiation; vulneryary;  
KW muscular degeneration; muscle formation; skeletal muscle differentiation;  
KW fat cell differentiation; degenerative muscle disease; muscle injury;  
KW gene; ds; chromosome 1p33.36.  
XX  
OS Homo sapiens.  
XX  
PN WO2003078463-A1.  
XX  
PD 25-SEP-2003.  
XX  
PF 13-MAR-2003; 2003WO-EP002638.  
XX  
PR 20-MAR-2002; 2002DE-01012397.  
XX  
PA (UNIV-) UNIVERSITAETSKLINIKUM FREIBURG.  
XX  
PI Schuele R, Hublitz P;  
XX  
DR WPI; 2003-767506/72.  
XX  
SQ New global repressor involved in myogenic differentiation protein, useful for treating disorders of skeletal muscle and fat cell differentiation, also related cDNA and antibodies.

Query Match 82.0%; Score 16.4; DB 10; Length 15105;  
Best Local Similarity 94.4%; Pred. No. 3.7e+02;

CC This invention describes a novel polypeptide that includes a sequence of at least 20 consecutive amino acids of human or murine GRIM1 (Global repressor involved in myogenic differentiation). The human GRIM1-encoding sequence is known, deposited as AL050019, and the corresponding murine sequence has been identified by comparative screening of murine expressed sequence tags (EST's). The products of the invention have vulneryary activity and the method of the invention can be used to modify GRIM function, by blocking movement of GRIM between nucleus and cytoplasm, to contract existing muscular degeneration or activate muscle formation. CC GRIM is a repressor of many complexes and promoters. The GRIM polypeptides are used to treat disorders of skeletal muscle differentiation or fat cell differentiation, including degenerative muscle diseases and muscle injuries in the elderly and to identify specific modulators. This sequence encodes the human GRIM1 protein. XX  
SQ Sequence 15105 BP; 2767 A; 3992 C; 4967 G; 3379 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 15105;  
Best Local Similarity 94.4%; Pred. No. 3.7e+02;

CC sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. CC The current sequence is that of a human soft tissue sarcoma-upregulated CC DNA of the invention. The current sequence is not shown within the CC specification per se but was submitted in CD format by the inventor. XX  
SQ Sequence 4560 BP; 803 A; 1537 C; 1370 G; 850 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 4560;  
Best Local Similarity 94.4%; Pred. No. 3.2e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20  
DB 2727 GCGCTCTCTGAGTAGCAG 2710  
|||||

RESULT 36  
ADQ97081/c  
ID ADQ97081 standard; DNA; 63411 BP.  
XX  
AC ADQ97081;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Mouse cancer associated sequence MD10-004, SEQ ID 57.  
XX  
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.  
XX  
OS Mus musculus.  
XX  
PN WO2004060304-A2.  
XX  
PD 22-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US041389.  
XX  
PR 27-DEC-2002; 2002US-00330773.  
XX  
PA (SAGR-) SAGRES DISCOVERY INC.  
XX  
PI Morris DW, Malandro MS;  
XX  
DR WPI; 2004-543781/52.  
XX  
PT New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.  
XX  
PS Claim 1; SEQ ID NO 57; 199pp; English.  
XX  
CC The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 63411 BP; 17408 A; 15472 C; 14344 G; 15847 T; 0 U; 340 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 63411;  
Best Local Similarity 94.4%; Pred. No. 4.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGC 18  
DB 29444 AGGGCTTCTCTGAGTAGC 29427  
|||||

RESULT 37  
ACH41376/c  
ID ACH41376 standard; cDNA; 458 BP.  
XX  
AC ACH41376;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human foetal brain cDNA #2743.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.





virucide; antibacterial; fungicide; cytostatic; neutropenia; leukopenia; chemotherapy; radiation therapy; HIV; AIDS; immunodeficiency disease; bacterial infection; acute myeloid leukaemia; antifungal therapy; systemic; invasive candidiasis; human; hg-CSF; ds; gene.

Homo sapiens.

Key Location/Qualifiers  
CDS 1..525  
/\*tag= a  
/partial  
/product= "Human granulocyte colony-stimulating factor protein"  
/note= "No start codon"

WO2003006501-A2.

23-JAN-2003.

10-JUL-2002; 2002WO-DK000482.

11-JUL-2001; 2001US-00904196.

22-MAR-2002; 2002DK-0000447.

08-MAY-2002; 2002DK-00000708.

(MAXY-) MAXYGEN HOLDINGS LTD.

Nissen TL, Andersen KV, Hansen CK, Mikkelsen JM, Schambye HT;

WPI; 2003-221717/21.

P-PSDB; ADK41219.

New polypeptide conjugate exhibiting granulocyte colony-stimulating factor activity, useful for preventing or treating neutropenia or leukopenia due to chemotherapy or radiation therapy, AIDS or other immunodeficiency diseases.

Example 1; SEQ ID NO 2; 106pp; English.

The invention relates to a novel polypeptide conjugate exhibiting granulocyte colony-stimulating factor activity (G-CSF). The polypeptide conjugate comprises: a polypeptide comprising at least one substitution selected from the group of K16R/Q, K34R/Q, and K40R/Q, and at least one substitution selected from the group of T105K and S159K relative to the amino acid sequence of hg-CSF having a 174 residue amino acid sequence, given in the specification, or in a corresponding position relative to an amino acid sequence having at least 80 % sequence identity with the 174 residue amino acid sequence, where the conjugate has 2-6 polyethylene glycol moieties with a molecular weight of 1000-10000 Da attached to attachment groups of the polypeptide; or a polypeptide comprising an amino acid sequence that differs in at least one amino acid residue from the amino acid sequence of hg-CSF having the sequence of the 174 residue amino acid sequence, and with at least one non-polypeptide moiety attached to an attachment group of the polypeptide, where the conjugate has an in vitro bioactivity in the range of 2-30% of the bioactivity of non-conjugated hg-CSF as determined by the luciferase assay. The G-CSF conjugate has immunostimulant, anti-HIV, virucide, antibacterial, fungicide, and cytostatic activities. The polypeptide conjugate is useful as a pharmaceutical, or for preparing a pharmaceutical composition for treating an insufficient neutrophil level. The pharmaceutical composition is also useful for preventing and/or treating neutropenia or leukopenia due to chemotherapy or radiation therapy, or due to HIV or another viral infection. The polypeptide conjugate may also be used for treating AIDS or other immunodeficiency diseases, bacterial infections, acute myeloid leukaemia, or for antifungal therapy, particularly for treating systemic or invasive candidiasis. This polynucleotide sequence represents the DNA encoding the human G-CSF protein of the invention.

Sequence 525 BP; 104 A; 143 C; 153 G; 125 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 10; Length 525;  
Best Local Similarity 89.5%; Pred. No. 5.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGCTCTCTGAGTAGCAG 20  
|||||  
DB 173 GGGCGCCTCTGAGTAGCTG 191

RESULT 40

ABS52517

ID ABS52517 standard; DNA; 531 BP.

XX ABS52517;

DT 15-NOV-2002 (first entry)

DE DNA encoding G-CSF copy 2 in the single chain G-CSF dimer.

Gene; ds; single-chain multimeric polypeptide; polyethylene glycol; PEG; granulocyte colony stimulating factor; G-CSF; human; haematopoietic disorder; radiation therapy; chemotherapy; bone marrow transplantation; acquired immunodeficiency syndrome; AIDS; immunodeficiency disease; leukopenia; acute myeloid leukaemia; half-life; clearance; immunogenicity; bioavailability; single chain G-CSF dimer; antihuman immunodeficiency virus; HIV; haemostatic.

OS Homo sapiens.

OS Synthetic.

Key Location/Qualifiers

CDS 1..525

/\*tag= a

/product= "G-CSF dimer"

/partial

/note= "This CDS encodes for residues 175-348 in the G-CSF dimer polypeptide. No start codon shown"

PN WO200236626-A1.

XX 10-MAY-2002.

XX 01-NOV-2001; 2001WO-DK000724.

XX 02-NOV-2000; 2000DK-00001647.

(MAXY-) MAXYGEN APS.

(MAXY-) MAXYGEN HOLDINGS LTD.

PI Nissen TL, Jensen AD;

WPI; 2002-618972/66.

P-PSDB; ABG32020.

Single chain multimeric polypeptide conjugate for treating hematopoietic disorders, has two units of monomeric polypeptides linked via peptide bond/linker and polymer group bound to attachment group of polypeptide.

Example 1; Page 94; 108pp; English.

The invention discloses a single-chain multimeric polypeptide conjugate comprising at least two units of a monomeric polypeptide linked via a peptide bond or a peptide linker, where the monomeric polypeptide is biologically active in its monomeric form and has at least one polymer group covalently bound to an attachment group of the polypeptide e.g. polyethylene glycol (PEG). The polypeptide is preferably a single-chain multimeric granulocyte colony stimulating factor (G-CSF) polypeptide comprising at least two G-CSF polypeptide monomers, linked via a peptide bond or a peptide linker, where at least one of the monomers is a variant of wild-type human G-CSF comprising at least one amino acid residue modification. The monomeric and multimeric polypeptides are useful in therapy and for manufacture of a medicament for treatment of general haematopoietic disorders, including disorders arising from radiation therapy, chemotherapy or bone marrow transplantations, acquired immunodeficiency syndrome (AIDS) or other immunodeficiency diseases, leukopenia and acute myeloid leukaemia. The conjugate has one or more

CC important properties as compared to the native polypeptide, including  
CC increased functional in vivo half-life, increased serum half-life,  
CC reduced clearance, reduced immunogenicity and/or increased  
CC bioavailability. Consequently, medical treatment with a conjugate offers  
CC advantages including longer duration between injections and fewer side  
CC effects. The sequence presented is the DNA encoding the G-CSF copy 2 in  
CC the single chain G-CSF dimer

XX SQ Sequence 531 BP; 106 A; 144 C; 154 G; 127 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 6; Length 531;  
Best Local Similarity 89.5%; Pred. No. 5.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
Db 173 GGGCGCTCTGAGTAGCTG 191

RESULT 41  
ABZ53961/c  
ID ABZ53961 standard; cDNA; 596 BP.

XX AC ABZ53961;  
XX DT 28-MAR-2003 (first entry)

XX DE Aspergillus oryzae polynucleotide SEQ ID NO 3074.

XX KW Aspergillus oryzae; fermentation; fungus; industrial; EST;  
XX KW expressed sequence tag; gene; ss.

XX OS Aspergillus oryzae.

XX PN WO200279476-A1.

XX PD 10-OCT-2002.

XX PF 22-MAR-2002; 2002WO-IB000890.

XX PR 30-MAR-2001; 2001JP-00098371.

XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PA (NORQ-) NAT RES INST BREWING.

XX PA (NORQ-) NAT FOOD RES INST MIN AGRIC.

XX PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;

XX PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;

XX DR WPI; 2003-046817/04.

XX PT Detection of expression of specific Aspergillus genes for monitoring the  
XX PT fermentation and growth conditions of the fungus, using DNA probes.

XX PS Claim 1; SEQ ID NO 3074; 48pp + Sequence Listing; Japanese.

XX CC The invention relates to a polynucleotide having any of 6006 specific  
XX CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under  
XX CC specific culture conditions including one or more of eutrophic,  
XX CC oligotrophic, solid, early germination, alkaline, high temperature, low  
XX CC temperature or maltose culture or polynucleotides stringently hybridising  
XX CC to these sequences. The polynucleotides are useful for monitoring the  
XX CC progress of fermentation and the growth conditions of a fungus,  
XX CC especially of Aspergillus oryzae which is widely used in industrial  
XX CC fermentation. Also monitoring for fungal contamination. Note: The  
XX CC sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 596 BP; 161 A; 153 C; 138 G; 144 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 8; Length 596;  
Best Local Similarity 89.5%; Pred. No. 5.2e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
Db 495 AGGGCGCTCTTAAGTAGCA 477

RESULT 42

AAA01703/c

ID AAA01703 standard; cDNA; 779 BP.

XX AC AAA01703;

XX DT 19-MAY-2000 (first entry)

XX DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:1694.

XX KW Human; colon cancer; tumour; diagnosis; gene expression product; probe;  
XX KW detection; cancerous state; metastasis; identification; breast cancer;  
XX KW oestrogen receptor-positive breast cancer; therapy;

XX KW oestrogen receptor-negative breast cancer; lung cancer; ss.

XX OS Homo sapiens.

XX PN WO9958675-A2.

XX PD 18-NOV-1999.

XX PF 13-MAY-1999; 99WO-US010602.

XX PR 14-MAY-1998; 98US-0085426P.

XX PR 15-MAY-1998; 98US-0085537P.

XX PR 15-MAY-1998; 98US-0085696P.

XX PR 21-OCT-1998; 98US-0105234P.

XX PR 27-OCT-1998; 98US-0105877P.

XX PA (CHIR) CHIRON CORP.

XX PA (HYSE-) HYSEQ INC.

XX PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
XX PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
XX PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
XX PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Grain B;

XX DR WPI; 2000-126369/11.

XX PT Polynucleotide library used to determine cancerous states of mammalian  
XX PT cells.

XX PS Claim 1; Page 616-617; 1097pp; English.

XX CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
XX CC libraries constructed from human colon cancer cell lines. The present  
XX CC invention also describes a method of detecting differentially expressed  
XX CC genes correlated with a cancerous state of a mammalian cell, comprising  
XX CC detecting at least one differentially expressed gene product in a test  
XX CC sample derived from a cell suspected of being cancerous, where detection  
XX CC of the differentially expressed gene product is correlated with a  
XX CC cancerous state of the cell from which the test sample was derived. The  
XX CC polynucleotide sequences can be used in a method for detecting  
XX CC differentially expressed genes correlated with a cancerous state of a  
XX CC mammalian cell. The polynucleotides can also be used as probes for  
XX CC detecting and mapping related genes. They can be used in diagnosis and  
XX CC prognosis of diseases and disorders (e.g. identification of pre-  
XX CC metastatic or metastatic cancerous states, stages of cancer, or  
XX CC responsiveness of cancer to therapy). This is particularly for breast  
XX CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
XX CC negative breast cancer, lung cancer, and colon cancer

XX SQ Sequence 779 BP; 183 A; 215 C; 186 G; 184 T; 0 U; 11 Other;

Query Match 79.0%; Score 15.8; DB 3; Length 779;  
Best Local Similarity 89.5%; Pred. No. 5.3e+02;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
 |||||  
 Db 341 GGGCGTCTCGGAGTAGGAG 323

RESULT 43  
 ABD12260/c  
 ID ABD12260 standard; DNA; 882 BP.  
 XX  
 AC ABD12260;  
 XX 29-JUL-2004 (first entry)  
 XX  
 DE Pseudomonas aeruginosa polynucleotide #10864.  
 XX  
 DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
 KW antibacterial.  
 KW  
 OS Pseudomonas aeruginosa.  
 XX  
 PN US6551795-B1.  
 XX  
 PD 22-APR-2003.  
 XX  
 PF 18-FEB-1999; 99US-00252991.  
 XX  
 PR 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX WPI; 2003-615309/58.  
 DR P-PSDB; ABO78689.  
 XX  
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX  
 PS Disclosure; SEQ ID NO 10864; 455pp; English.  
 XX  
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biotech technology. Sequences ABD01397-  
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 882 BP; 208 A; 299 C; 231 G; 144 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 11; Length 882;  
 Best Local Similarity 89.5%; Pred. No. 5.4e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
 |||||  
 Db 177 GGGCGTCTCGGAGTAGCAG 159

RESULT 44

ABD12216  
 ID ABD12216 standard; DNA; 1203 BP.  
 XX  
 AC ABD12216;  
 XX 29-JUL-2004 (first entry)  
 XX  
 DE Pseudomonas aeruginosa polynucleotide #10820.  
 XX  
 DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;  
 KW antibacterial.  
 KW  
 OS Pseudomonas aeruginosa.  
 XX  
 PN US6551795-B1.  
 XX  
 PD 22-APR-2003.  
 XX  
 PF 18-FEB-1999; 99US-00252991.  
 XX  
 PR 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX WPI; 2003-615309/58.  
 DR P-PSDB; ABO78645.  
 XX  
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX  
 PS Disclosure; SEQ ID NO 10820; 455pp; English.  
 XX  
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biotech technology. Sequences ABD01397-  
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 1203 BP; 210 A; 320 C; 396 G; 277 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 11; Length 1203;  
 Best Local Similarity 89.5%; Pred. No. 5.6e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
 |||||  
 Db 760 GGGCGTCTCGGAGTAGCAG 778

RESULT 45  
 ABD07772  
 ID ABD07772 standard; cDNA; 2069 BP.  
 XX  
 AC ABD07772;  
 XX 04-NOV-2004 (first entry)  
 XX  
 DE Full length human cDNA useful for treating neurological disease Seq 1278.

XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;  
 KW osteoporosis; neurological disease; Alzheimer's disease;  
 KW Parkinson's disease; dementia; short memory; cancer;  
 KW sense or motor function; emotional reaction; fear response; panic;  
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytotatic;  
 KW tranquiliser.  
 XX Homo sapiens.  
 XX EP1447413-A2.  
 XX 18-AUG-2004.  
 XX 12-FEB-2004; 2004EP-000031145.  
 XX 14-FEB-2003; 2003JP-00102207.  
 XX 09-MAY-2003; 2003JP-00131452.  
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;  
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;  
 XX WPI: 2004-583265/57.  
 XX P-PSDB; ADR09728.  
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,  
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
 XX Claim 1; SEQ ID NO 1278; 2686pp; English.  
 XX This invention relates to novel, isolated full length human cDNA  
 CC molecules and the encoded proteins thereof. Specifically, it refers to  
 CC cDNA clones obtained by an oligo-capping method, where none of these  
 CC clones are identical to any known human mRNAs. The present invention  
 CC describes an immunoassay to identify agonists and antagonists, as well as  
 CC antibodies, antisense molecules and siRNAs that can all be used to bind  
 CC to and modulate expression of the cDNA molecules. As such, these  
 CC molecules are useful for diagnostic markers or therapeutic targets for  
 CC the various diseases or morbid states. In particular, they are useful in  
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's  
 CC disease, Parkinson's disease, dementia, short memory and various cancers,  
 CC as well as for maintaining equilibrium of sense or motor function, and  
 CC for treating emotional reaction, fear response and panic. Accordingly,  
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,  
 CC cytotatic and tranquiliser activities. This polynucleotide is a full  
 CC length human cDNA sequence of the invention. NOTE: This sequence is not  
 CC given in the sequence listing of the specification but can be obtained on  
 CC CD-ROM from the European Patent Office, Vienna Sub-office.  
 XX CD-ROM from the European Patent Office, Vienna Sub-office.  
 XX Sequence 2069 BP; 498 A; 540 C; 586 G; 445 T; 0 U; 0 Other;  
 SQ Query Match 79.0%; Score 15.8; DB 13; Length 2069;  
 Best Local Similarity 89.5%; Pred. No. 5.9e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GGGCGTCTCTGAGTAGCAG 20  
 DB 573 GGGCGTTTCTGAGGAGCAG 591  
 RESULT 46  
 AAV48219/c  
 ID AAV48219 standard; DNA; 2614 BP.  
 XX AAV48219;  
 AC AAV48219;  
 XX 09-NOV-1998 (first entry)  
 DT 09-NOV-1998 (first entry)  
 DE Mouse striated muscle preferentially expressed gene.  
 XX ds; human; striated muscle preferentially expressed protein;  
 KW

KW smooth muscle; cell proliferation; developmental stage; gamma-interferon;  
 KW tissue plasminogen activator; p21 cell cycle; nitric oxide synthetase.  
 OS Mus sp.  
 FH Key Location/Qualifiers  
 CDS 1..1803  
 FT /\*tag= a  
 FT /product= "Striated muscle preferentially expressed  
 FT protein"  
 FT /note= "No stop codon given in protein, however stop  
 FT codon present at position 1807..1809"  
 XX WO9835040-A2.  
 XX 13-AUG-1998.  
 XX 06-FEB-1998; 98WO-US002441.  
 XX 06-FEB-1997; 97US-00795868.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Lee M, Hsieh C;  
 PI WPI: 1998-447237/38.  
 DR P-PSDB; AAW77049.  
 XX Novel human, rat or mouse aorta or striated-muscle preferentially  
 PT expressed proteins - useful for treating e.g. atherosclerosis.  
 XX Claim 18; Fig 22; 88pp; English.  
 XX The striated muscle preferentially expressed protein (SPEG) can be linked  
 CC to an enhancer/promoter derived from an aortic preferentially expressed  
 CC gene to regulate smooth muscle cell-specific expression. This can be used  
 CC as a method of inhibiting vascular smooth muscle cell proliferation. The  
 CC nucleic acids are used to direct developmental stage-specific expression  
 CC of a heterologous polypeptide which is especially selected from tissue  
 CC plasminogen activator (tPA), p21 cell cycle inhibitor, nitric oxide  
 CC synthetase, gamma-interferon, atrial natriuretic proteins. These are used  
 CC to inhibit the proliferation of smooth muscle cells, e.g. for the  
 CC treatment of atherosclerosis  
 XX Sequence 2614 BP; 540 A; 816 C; 811 G; 447 T; 0 U; 0 Other;  
 SQ Query Match 79.0%; Score 15.8; DB 2; Length 2614;  
 Best Local Similarity 89.5%; Pred. No. 6.1e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GGGCGTCTCTGAGTAGCAG 20  
 DB 1273 GGGCGTCTCTGAGGAGCAG 1255  
 RESULT 47  
 AAZ51045/c  
 ID AAZ51045 standard; cDNA; 2614 BP.  
 XX AAZ51045;  
 AC AAZ51045;  
 XX 05-JUN-2000 (first entry)  
 DT 05-JUN-2000 (first entry)  
 DE Mouse striated muscle preferentially expressed gene partial cDNA.  
 XX Mouse; striated muscle preferentially expressed gene; SPEG;  
 KW aortic-preferentially-expressed gene-1; APEG-1; striated muscle cell;  
 KW aorta; smooth muscle cell; antiarteriosclerotic; vasotropic; treatment;  
 KW diagnosis; vascular disease; atherosclerosis; restenosis; ss.  
 XX Mus sp.  
 OS Mus sp.  
 FH Key Location/Qualifiers



CC therapeutic agents or belonging to risk groups (e.g. UV-mediated skin  
 CC damage, skin cancer and cancers associated with oxidative stress. The  
 CC method allows the characterisation and mechanism of the GSTT1 deletion  
 CC and identifies 18 kb homology regions flanking GSTT1 which are involved  
 CC in the deletion event that produced the \*0 allele. The method permits the  
 CC unambiguous discrimination of all GSTT1 genotypes (\*A/A, \*0/0 (both  
 CC homozygous), \*A/0 (heterozygous)). The technique allows the reproducible  
 CC simultaneous discrimination of all the genotypes. The three GSTT1  
 CC genotypes detected by these procedure correlated highly significant with  
 CC enzyme activity in erythrocytes. The trimodular distribution of  
 CC phenotypes at high-, intermediate- and null- activity in homo- and  
 CC heterozygotes for the \*A allele and \*0/0 homozygotes, respectively  
 CC indicate a gene dose effect. The present sequence is 5kb of genomic  
 CC sequence, 5' to the deletion point, from which PCR primers were designed  
 CC enabling the three alleles to be discriminated in an assay of the  
 CC invention  
 XX  
 SQ Sequence 5000 BP; 1230 A; 1231 C; 1464 G; 1075 T; 0 U; 0 Other;  
 Query Match 79.0%; Score 15.8; DB 4; Length 5000;  
 Best Local Similarity 89.5%; Pred. No. 6.5e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GGGCGTCTCTGAGTAGCAG 20  
 DB 3623 GGGCGTCTCTGAGTAGCAG 3605  
 AC ADK40991;  
 DT 06-MAY-2004 (first entry)  
 XX Novel human kinase gene #11.  
 DE  
 XX cytostatic; immunomodulator; cardiant; neuroprotective; nootropic;  
 KW antiparkinsonian; virucide; antibacterial; fungicide; ophthalmological;  
 KW analgesic; hypotensive; immunosuppressive; kinase inhibitor; kinase;  
 KW cancer; peripheral nervous system; central nervous system;  
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; viral infection; prion infection;  
 KW ocular disease; migraine; pain; sexual dysfunction; mood disorder;  
 KW attention disorder; cognition disorder; hypotension; hypertension;  
 KW psychotic disorder; neurological disorder; dyskinesia;  
 KW metabolic disorder; organ transplant rejection; enzyme; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003057841-A2.  
 XX  
 PD 17-JUL-2003.  
 XX  
 XX 31-DEC-2002; 2002WO-US041687.  
 PF  
 XX 31-DEC-2001; 2001US-0343169P.  
 PR  
 XX (GRIG/) GRIGORIEV I V.  
 PA (SUDA/) SUDARSANAM S.  
 XX  
 XX Grigoriev IV, Sudarsanam S;  
 PI WPI; 2003-587115/55.  
 DR  
 XX New isolated, enriched or purified nucleic acid molecule encoding a  
 PT kinase polypeptide, useful for treating cancer, immune-related diseases,  
 PT cardiovascular disease, brain or neuronal-associated diseases and  
 PT metabolic disorders.  
 XX  
 PS Claim 33; SEQ ID NO 98; 491pp; English.  
 XX

CC The invention relates to novel isolated, enriched or purified nucleic acid  
 CC molecules encoding a kinase polypeptide. The nucleic acid molecule  
 CC comprises a sequence that: (a) encodes a kinase polypeptide; (b) is a  
 CC complement of (a); (c) hybridizes under stringent conditions to (a) and  
 CC encodes a naturally occurring kinase polypeptide; (d) encodes the  
 CC polypeptide in (a), except that it lacks one or more, but not all, of an  
 CC N-terminal domain, C-terminal catalytic domain, a spacer region and a C-  
 CC terminal domain, a coiled-coil structure region, a catalytic domain, a C-  
 CC terminal tail; or (e) is a complement of (d). The nucleic acid molecules,  
 CC polypeptides, methods and substance are useful for treating cancers,  
 CC immune-related diseases or disorders, cardiovascular disease, brain or  
 CC neuronal-associated diseases, and metabolic disorders. The disorders are  
 CC preferably cancers of the tissues or of hematopoietic origin, diseases of  
 CC the central or peripheral nervous system, Alzheimer's disease,  
 CC Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis,  
 CC viral infections, infections caused by prions, infections caused by  
 CC bacterial, infections caused by fungi, ocular diseases, migraines, pain,  
 CC sexual dysfunction, mood disorders, attention disorders, cognition  
 CC disorders, hypotension, hypertension, psychotic disorders, neurological  
 CC disorders, dyskinesias, metabolic disorders and organ transplant  
 CC rejection. This sequence corresponds to the DNA encoding one of the  
 CC kinase polypeptides of the invention.  
 XX

SQ Sequence 5727 BP; 1464 A; 1397 C; 1637 G; 1229 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 10; Length 5727;

Best Local Similarity 89.5%; Pred. No. 6.6e+02; Length 5727;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20

|||||

DB 3911 GGGCGTCTCTGAGTAGGAG 3893

Search completed: March 14, 2005, 20:46:06  
 Job time : 294 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 17:14:07 ; Search time 1891 Seconds  
(without alignments)  
402.583 Million cell updates/sec

Title: US-10-672-399-10

Perfect score: 20

Sequence: 1 agggcgctctctgagtagcag 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	232	6	CB266804
C 2	20	100.0	551	4	BG707429
C 3	20	100.0	551	4	BG707608
C 4	20	100.0	610	5	BX327795
C 5	20	100.0	739	9	AY421570 Pan trogl
C 6	20	100.0	782	4	BI753116
C 7	20	100.0	896	5	BX371636 BX371636
C 8	20	100.0	1084	4	BM544718 AGENCOURT
C 9	20	100.0	1728	9	AY421569
C 10	20	100.0	2072	3	CR602106
C 11	18	90.0	1087	1	AL571167
C 12	17.4	87.0	385	8	BH235764 ATZKC34TF
C 13	17.4	87.0	521	8	AZ408342
C 14	17.4	87.0	578	7	CO708502
C 15	17.4	87.0	593	8	BH236262
C 16	17.4	87.0	687	9	CE013803
C 17	17.4	87.0	1746	9	AY421571 Mus muscu
C 18	16.8	84.0	210	8	AZ850382
C 19	16.8	84.0	249	9	CR037148
C 20	16.8	84.0	252	9	CE277790
C 21	16.8	84.0	402	7	R2012636
C 22	16.8	84.0	464	7	R23689
C 23	16.8	84.0	519	8	AQ229055
C 24	16.8	84.0	547	9	CR265831

C 25	16.8	84.0	550	9	CE015183
C 26	16.8	84.0	551	7	CV458871
C 27	16.8	84.0	554	4	BI788920
C 28	16.8	84.0	565	6	CA248736
C 29	16.8	84.0	569	4	BM335950
C 30	16.8	84.0	593	9	CG985592
C 31	16.8	84.0	645	8	BH646059
C 32	16.8	84.0	647	8	AZ966847
C 33	16.8	84.0	664	8	AZ840734
C 34	16.8	84.0	669	7	CV288705
C 35	16.8	84.0	679	9	CR191252
C 36	16.8	84.0	684	1	AV370109
C 37	16.8	84.0	702	9	AG119191
C 38	16.8	84.0	708	7	CV289631
C 39	16.8	84.0	723	7	CV289106
C 40	16.8	84.0	730	8	BH316461
C 41	16.8	84.0	736	9	AG383658
C 42	16.8	84.0	744	9	AG569158
C 43	16.8	84.0	744	9	CC552929
C 44	16.8	84.0	764	9	BX982971
C 45	16.8	84.0	773	9	AG591654
C 46	16.8	84.0	773	9	CV292304
C 47	16.8	84.0	796	9	CL662911
C 48	16.8	84.0	846	7	CV289382
C 49	16.8	84.0	855	9	CR802378
C 50	16.8	84.0	874	9	CR194037
C 51	16.8	84.0	876	7	CV291568
C 52	16.8	84.0	920	7	CV289351
C 53	16.8	84.0	962	9	CNS038EA
C 54	16.8	84.0	1035	9	CNS068SUH
C 55	16.8	84.0	1407	3	AK033454
C 56	16.8	84.0	2800	3	AK051660
C 57	16.8	84.0	2931	3	AK035662
C 58	16.4	82.0	166	9	CE771307
C 59	16.4	82.0	201	2	BF734506
C 60	16.4	82.0	272	2	BF907757
C 61	16.4	82.0	283	8	AZ492804
C 62	16.4	82.0	303	2	BF394207
C 63	16.4	82.0	327	7	TI9249
C 64	16.4	82.0	406	6	CB142104
C 65	16.4	82.0	446	4	BM827484
C 66	16.4	82.0	474	2	BF393450
C 67	16.4	82.0	484	2	BF394808
C 68	16.4	82.0	499	2	AW962815
C 69	16.4	82.0	514	9	AG250246
C 70	16.4	82.0	538	2	BE742897
C 71	16.4	82.0	545	6	CA741357
C 72	16.4	82.0	552	7	CN300697
C 73	16.4	82.0	554	6	CB131572
C 74	16.4	82.0	568	4	BI254811
C 75	16.4	82.0	572	2	BE296203
C 76	16.4	82.0	581	5	BP366762
C 77	16.4	82.0	585	1	AL046149
C 78	16.4	82.0	587	4	BG436830
C 79	16.4	82.0	587	6	CA504333
C 80	16.4	82.0	588	7	CK889521
C 81	16.4	82.0	595	8	AZ828093
C 82	16.4	82.0	599	8	AZ710802
C 83	16.4	82.0	607	5	BP245334
C 84	16.4	82.0	630	7	CB430372
C 85	16.4	82.0	643	4	BJ262582
C 86	16.4	82.0	644	7	CN300707
C 87	16.4	82.0	651	4	BG761400
C 88	16.4	82.0	671	2	BE779713
C 89	16.4	82.0	673	4	BM042146
C 90	16.4	82.0	682	4	BJ257007
C 91	16.4	82.0	689	2	BE273057
C 92	16.4	82.0	699	7	CN300693
C 93	16.4	82.0	700	4	BG388020
C 94	16.4	82.0	707	4	BM006304
C 95	16.4	82.0	730	6	CB499497
C 96	16.4	82.0	740	4	BG281880
C 97	16.4	82.0	743	1	AI066022

CE015183	tigr-gss-
CV458871	aofo1-9ms
BI788920	ie41c05.Y
CA248736	SCSBFL110
BM335950	MEST179-D
CG985592	CH240.155
BH646059	BOH2N67TF
AZ966847	2M0237106
AZ840734	2M0138122
CV288705	aofo1-9ms
CR191252	Reverse s
AV370109	AV370109
AG119191	Pan trogl
CV289631	aofo1-8ms
CV289106	aofo1-1ms
BH316461	CH230-40G
AG383658	Mus muscu
AG569158	Mus muscu
CC552929	CH240.438
AG591654	Mus muscu
CV292304	aofo1-11m
CL662911	PR10142d
CV289382	aofo1-2ms
CR802378	GR0AAA21D
CR194037	Reverse s
CV291568	aofo1-16m
CV289351	aofo1-7ms
AL232507	Tetraodon
AL413727	T7 end of
AK033454	Mus muscu
AK051660	Mus muscu
AK035662	Mus muscu
CE771307	tigr-gss-
BF734506	PM2-AN008
BF907757	IL2-UT007
AZ492804	IM0327D07
BF394207	UI-R-CAO-
TI9249	d03008T Tes
CB142104	K-EST0195
BM827484	K-EST0099
BF393450	UI-R-CAO-
BF394808	UI-R-CAO-
AW962815	EST374888
AG250246	Lotus cor
BE742897	601574583
CA741357	wia1c.pk0
CN300697	170004177
CB131572	K-EST0181
BI254811	602974057
BE296203	601177354
BP366762	BP366762
AL046149	DKP2P434A
BG436830	602488817
CA504333	UI-R-FJO-
CK889521	SGP161545
AZ828093	2M0104G24
AZ710802	RPCI-24-1
BP245334	BP245334
CB430372	OJ48e11.Y
BJ262582	BJ262582
CN300707	170006001
BG761400	602718289
BE779713	601465250
BM042146	603615783
BJ257007	BJ257007
BE273057	601171394
CN300693	170005999
BG388020	602412861
BM006304	603614984
CB499497	88a1hnc01
BG281880	602402619
AI066022	B8BmMFS20





244	15.8	79.0	495	8	BH070793	BH070793	RPCI-24-3	317	15.8	79.0	720	8	A2825641	A2825641	2M0100N21
245	15.8	79.0	499	1	AI129469	QC44D12.X	CE446925	318	15.8	79.0	724	2	BF430769	BF430769	OG04B02T3
246	15.8	79.0	518	8	CA466925	tigr-g88-	CE466925	319	15.8	79.0	725	6	CA259298	CA259298	SCEQRT301
247	15.8	79.0	521	2	BE097188	UI-R-B01-	BE097188	320	15.8	79.0	728	8	AK233097	AK233097	IM0056N05
248	15.8	79.0	524	4	BG789955	saee1b09.	BG789955	321	15.8	79.0	730	7	CK947974	CK947974	4072678 B
249	15.8	79.0	525	1	AA848303	EST191063	AA848303	322	15.8	79.0	730	9	AG390119	AG390119	Mus muscu
250	15.8	79.0	525	1	AI579376	UI-R-Y0-V	AI579376	323	15.8	79.0	733	6	CA505500	CA505500	UI-R-FS1-
251	15.8	79.0	531	8	AZ365345	IM0111024	AZ365345	324	15.8	79.0	733	6	BI689926	BI689926	603310760
252	15.8	79.0	535	6	CB148110	K-EST0204	CB148110	325	15.8	79.0	739	9	AG374036	AG374036	Mus muscu
253	15.8	79.0	538	6	CA228994	SCQGPL305	CA228994	326	15.8	79.0	740	5	BQ204580	BQ204580	UI-R-D21-
254	15.8	79.0	543	2	BF403656	UI-R-CA1-	BF403656	327	15.8	79.0	747	6	CF069084	CF069084	EST669805
255	15.8	79.0	544	9	CE504906	tigr-g88-	CE504906	328	15.8	79.0	747	7	CR452535	CR452535	CR452535
256	15.8	79.0	544	9	CL368014	RPCI44-37	CL368014	329	15.8	79.0	748	7	CO396018	CO396018	AGENCOURT
257	15.8	79.0	545	6	CA142717	SCMCR210	CA142717	330	15.8	79.0	749	3	CR714545	CR714545	Tetraodon
258	15.8	79.0	549	2	BF321524	uz62905.Y	BF321524	331	15.8	79.0	751	6	CA919645	CA919645	EST637363
259	15.8	79.0	551	2	BE232818	138051.MA	BE232818	332	15.8	79.0	767	8	AZ664038	AZ664038	IM0544F01
260	15.8	79.0	567	2	BE755515	209379.NA	BE755515	333	15.8	79.0	772	5	BQ208145	BQ208145	UI-R-EP0-
261	15.8	79.0	570	7	CF137591	UI-HP-BN0	CF137591	334	15.8	79.0	772	9	AG501097	AG501097	Mus muscu
262	15.8	79.0	570	9	CE073212	tigr-g88-	CE073212	335	15.8	79.0	779	6	CA157161	CA157161	SCEQRZ302
263	15.8	79.0	572	5	BQ208206	UI-R-EP0-	BQ208206	336	15.8	79.0	780	7	CF872648	CF872648	trIC002XJ
264	15.8	79.0	572	8	BZ894473	Hg13_0123	BZ894473	337	15.8	79.0	781	4	BG911931	BG911931	602809627
265	15.8	79.0	579	6	CD886512	GL18_1021	CD886512	338	15.8	79.0	796	6	CB667673	CB667673	OSJNEd15B
266	15.8	79.0	585	8	BZ561288	pac82-164	BZ561288	339	15.8	79.0	797	6	CB565688	CB565688	OSJNEc11L
267	15.8	79.0	586	2	BF112935	EST440618	BF112935	340	15.8	79.0	797	7	CO811183	CO811183	AGENCOURT
268	15.8	79.0	588	4	BI806570	S068E01.S	BI806570	341	15.8	79.0	805	8	BZ577010	BZ577010	meH2_5208
269	15.8	79.0	588	4	BM085976	sah33h09.	BM085976	342	15.8	79.0	806	6	CB322408	CB322408	UI-R-DY0-
270	15.8	79.0	589	8	AQ614855	HS_5140.A	AQ614855	343	15.8	79.0	814	8	BZ574508	BZ574508	meH2_3703
271	15.8	79.0	594	5	BQ208213	UI-R-EP0-	BQ208213	344	15.8	79.0	815	4	BI183098	BI183098	UNL-P-FN-
272	15.8	79.0	594	7	CR165714	Reverse.8	CR165714	345	15.8	79.0	815	8	BZ574646	BZ574646	meH2_3766
273	15.8	79.0	599	7	CO134881	EST829552	CO134881	346	15.8	79.0	821	4	BI182173	BI182173	UNL-P-FN-
274	15.8	79.0	600	5	BQ570796	UI-M-FB0-	BQ570796	347	15.8	79.0	821	6	CB322420	CB322420	UI-R-DY0-
275	15.8	79.0	600	6	CA757620	OE04B12-T	CA757620	348	15.8	79.0	823	6	CB669819	CB669819	OSJNEe02J
276	15.8	79.0	601	1	AU082994	AU082994	AU082994	349	15.8	79.0	830	7	CV139223	CV139223	EST850432
277	15.8	79.0	603	1	AL703709	DKF2p686D	AL703709	350	15.8	79.0	836	7	CN809455	CN809455	EST0801.M
278	15.8	79.0	606	5	BQ211501	UI-R-DY1-	BQ211501	351	15.8	79.0	850	6	CA131389	CA131389	SCEPRT107
279	15.8	79.0	611	9	CE309930	tigr-g88-	CE309930	352	15.8	79.0	850	6	CD758183	CD758183	AGENCOURT
280	15.8	79.0	612	4	BI309184	EST30594	BI309184	353	15.8	79.0	850	7	CO235393	CO235393	WS0264.B2
281	15.8	79.0	612	7	CF670348	RTCNT1_49	CF670348	354	15.8	79.0	853	4	BI183135	BI183135	UNL-P-FN-
282	15.8	79.0	614	2	BB658807	BB658807	BB658807	355	15.8	79.0	858	6	CB635977	CB635977	OSIIEB16P
283	15.8	79.0	622	2	BB619137	BB619137	BB619137	356	15.8	79.0	858	7	CK140982	CK140982	AGENCOURT
284	15.8	79.0	622	6	C77492	C77492	C77492	357	15.8	79.0	860	5	EX928485	EX928485	EX928485
285	15.8	79.0	624	4	BI088236	602851292	BI088236	358	15.8	79.0	861	9	CG870336	CG870336	ZMMBHC027
286	15.8	79.0	628	6	CA287018	SCSBSD205	CA287018	359	15.8	79.0	862	7	CO213769	CO213769	WS00930.B
287	15.8	79.0	634	8	AZ725179	RPCI-24-7	AZ725179	360	15.8	79.0	863	9	CR103638	CR103638	Reverse.8
288	15.8	79.0	634	9	CL714494	OR_Bba003	CL714494	361	15.8	79.0	882	5	BU535451	BU535451	AGENCOURT
289	15.8	79.0	641	1	AI133192	HA1787.Hu	AI133192	362	15.8	79.0	884	7	CK770217	CK770217	958358.MA
290	15.8	79.0	642	9	CL342992	RPCI44_27	CL342992	363	15.8	79.0	900	5	BQ644902	BQ644902	AGENCOURT
291	15.8	79.0	648	7	CF670268	RTCNT1_49	CF670268	364	15.8	79.0	907	4	BI252189	BI252189	602952882
292	15.8	79.0	653	6	CA237780	SCACFL502	CA237780	365	15.8	79.0	907	8	CC092538	CC092538	CSU-K34.1
293	15.8	79.0	653	8	AZ107821	RPCI-23-3	AZ107821	366	15.8	79.0	913	5	BQ919821	BQ919821	AGENCOURT
294	15.8	79.0	655	9	CL736005	OR_Bba006	CL736005	367	15.8	79.0	935	5	BU600244	BU600244	AGENCOURT
295	15.8	79.0	657	9	CE665421	tigr-g88-	CE665421	368	15.8	79.0	937	2	BE573380	BE573380	601333046
296	15.8	79.0	658	2	BB638418	BB638418	BB638418	369	15.8	79.0	953	9	CO099103	CO099103	ISB1-3202
297	15.8	79.0	658	6	CD917117	G608_1048	CD917117	370	15.8	79.0	983	5	BQ652867	BQ652867	AGENCOURT
298	15.8	79.0	663	4	BG459556	NF070F10P	BG459556	371	15.8	79.0	986	4	BG252608	BG252608	602365242
299	15.8	79.0	664	9	CE197328	tigr-g88-	CE197328	372	15.8	79.0	1006	2	BB609618	BB609618	BB609618
300	15.8	79.0	667	6	CA144711	SCSFR206	CA144711	373	15.8	79.0	1046	9	AG103111	AG103111	Pan trogl
301	15.8	79.0	675	6	CA153474	SCUTR2202	CA153474	374	15.8	79.0	1059	8	BZ550919	BZ550919	Pac81-60
302	15.8	79.0	675	8	AZ881605	RPCI-23-2	AZ881605	375	15.8	79.0	1060	4	BM551157	BM551157	AGENCOURT
303	15.8	79.0	676	6	CD768857	AGENCOURT	CD768857	376	15.8	79.0	1061	5	BQ943231	BQ943231	AGENCOURT
304	15.8	79.0	676	8	BZ879838	CH240_293	BZ879838	377	15.8	79.0	1075	4	BM549499	BM549499	AGENCOURT
305	15.8	79.0	676	9	AG142184	Pan trogl	AG142184	378	15.8	79.0	1120	3	AX009264	AX009264	Mus muscu
306	15.8	79.0	681	2	BB125916	BB125916	BB125916	379	15.8	79.0	1131	2	BF688484	BF688484	602185089
307	15.8	79.0	683	6	CA189499	SCCCLRIC0	CA189499	380	15.8	79.0	1170	3	AP116668	AP116668	Homo sapi
308	15.8	79.0	685	4	BJ167394	BJ167394	BJ167394	381	15.8	79.0	1173	8	CC211181	CC211181	CH261-186
309	15.8	79.0	687	6	CA126757	SCVPLR100	CA126757	382	15.8	79.0	1248	9	CL945843	CL945843	ObiFSB000
310	15.8	79.0	690	2	BB628254	BB628254	BB628254	383	15.8	79.0	1605	3	AK012974	AK012974	Mus muscu
311	15.8	79.0	698	9	AG361797	Mus muscu	AG361797	384	15.8	79.0	1628	9	AG039726	AG039726	Pan trogl
312	15.8	79.0	701	6	CB565697	OSJNEc11L	CB565697	385	15.8	79.0	2348	3	AK031142	AK031142	Mus muscu
313	15.8	79.0	702	3	CR594966	CR594966	CR594966	386	15.8	79.0	2363	3	AK033189	AK033189	Mus muscu
314	15.8	79.0	702	7	CK967262	Tetraodon	CK967262	387	15.8	79.0	2445	3	AK032530	AK032530	Mus muscu
315	15.8	79.0	709	9	AG566979	Mus muscu	AG566979	388	15.8	79.0	2504	3	AK036778	AK036778	Mus muscu
316	15.8	79.0	712	9	AG336188	Mus muscu	AG336188	389	15.8	79.0	2522	3	BC042506	BC042506	Mus muscu

C 390	15.8	79.0	2549	3	AK041713	Mus muscu
C 391	15.8	79.0	2559	3	AK084516	Mus muscu
C 392	15.8	79.0	2943	3	AK016569	Mus muscu
C 393	15.8	79.0	2999	3	AK087790	Mus muscu
C 394	15.8	79.0	3375	3	AK085739	Mus muscu
C 395	15.8	79.0	3479	3	AK035841	Mus muscu
C 396	15.8	79.0	6616	3	AK090124	Mus muscu
C 397	15.4	77.0	141	1	AV012267	AV012267
C 398	15.4	77.0	279	2	BB358001	BB358001
C 399	15.4	77.0	313	1	AA355948	EST64473
C 400	15.4	77.0	335	2	BF326376	PN3-AN009
C 401	15.4	77.0	341	8	AQ065272	HS 2179_B
C 402	15.4	77.0	344	7	CN955224	2721_52-3
C 403	15.4	77.0	397	6	CA194388	SCRFSB102
C 404	15.4	77.0	406	5	BY227135	BY227135
C 405	15.4	77.0	423	1	AA584698	no09G12.8
C 406	15.4	77.0	442	7	CN956589	4089_5001
C 407	15.4	77.0	445	8	AQ057771	CIT-HSP-2
C 408	15.4	77.0	455	9	CR158349	Reverse s
C 409	15.4	77.0	462	4	BG471874	602513351
C 410	15.4	77.0	487	1	AT1716236	UI-R-YO-a
C 411	15.4	77.0	489	6	CD122121	MEI-0071P
C 412	15.4	77.0	499	7	CF608675	GERMO01_0
C 413	15.4	77.0	515	8	AQ398204	mgxb0010E
C 414	15.4	77.0	521	6	CD922380	G750_102J
C 415	15.4	77.0	540	8	AZ097714	RPCI-23-4
C 416	15.4	77.0	542	5	BP227693	BP227693
C 417	15.4	77.0	572	5	BP340000	BP340000
C 418	15.4	77.0	581	5	BP254330	BP254330
C 419	15.4	77.0	583	5	BP312166	BP312166
C 420	15.4	77.0	584	8	AZ600021	1M0416I08
C 421	15.4	77.0	585	5	BP276512	BP276512
C 422	15.4	77.0	585	6	CD911442	G550_111C
C 423	15.4	77.0	600	6	CA060106	esaltfGB51
C 424	15.4	77.0	605	4	BM737043	952052F09
C 425	15.4	77.0	611	7	CJ023475	426_Full1
C 426	15.4	77.0	618	4	RV298299	RV298299
C 427	15.4	77.0	630	5	BQ063382	WHE3578_B
C 428	15.4	77.0	635	7	CN982624	51114_126
C 429	15.4	77.0	637	8	AZ493972	1M0328L23
C 430	15.4	77.0	638	4	BG330135	602429821
C 431	15.4	77.0	639	4	BG324506	602422455
C 432	15.4	77.0	652	7	CF660950	CcLL10a33
C 433	15.4	77.0	657	4	BG325318	602423893
C 434	15.4	77.0	665	7	CN984018	52734_126
C 435	15.4	77.0	669	5	BX861635	BX861635
C 436	15.4	77.0	677	2	BF315913	601895851
C 437	15.4	77.0	687	2	BE298970	601119502
C 438	15.4	77.0	688	6	CD889337	G118_111K
C 439	15.4	77.0	691	4	BM280436	D2B4_Di6C
C 440	15.4	77.0	691	9	CL384701	RPC144_32
C 441	15.4	77.0	695	6	CA059861	esaltfGB53
C 442	15.4	77.0	697	2	BE262251	601153069
C 443	15.4	77.0	713	2	BE294714	601173938
C 444	15.4	77.0	721	9	CL384721	RPC144_32
C 445	15.4	77.0	723	4	BG170793	602323725
C 446	15.4	77.0	723	4	BI261567	602933647
C 447	15.4	77.0	730	8	BH036802	RPCI-24-3
C 448	15.4	77.0	741	9	CG295278	OGW1047TH
C 449	15.4	77.0	742	6	CA431768	UI-H-DF0-
C 450	15.4	77.0	754	1	AL706268	DFZp6860
C 451	15.4	77.0	760	6	CA489510	AGENCOURT
C 452	15.4	77.0	764	4	BG930069	602764811
C 453	15.4	77.0	781	4	BI145297	602910419
C 454	15.4	77.0	786	4	BG685473	602637335
C 455	15.4	77.0	789	4	BI833116	603090832
C 456	15.4	77.0	794	4	BI907560	603065920
C 457	15.4	77.0	796	8	BZ532518	OGA1Y03TC
C 458	15.4	77.0	800	5	EX880681	EX880681
C 459	15.4	77.0	803	2	BE729095	601561139
C 460	15.4	77.0	809	4	BI755598	603027295
C 461	15.4	77.0	812	7	CF662513	CcLL02a06
C 462	15.4	77.0	813	4	BI260257	602972830

C 463	15.4	77.0	815	4	BG386170	BG386170
C 464	15.4	77.0	815	8	B2107586	CH230-134
C 465	15.4	77.0	827	2	BE561700	601345766
C 466	15.4	77.0	827	5	BQ710678	AGENCOURT
C 467	15.4	77.0	858	4	BI911176	603062756
C 468	15.4	77.0	862	4	BI261265	602972930
C 469	15.4	77.0	868	4	BG759673	602713366
C 470	15.4	77.0	874	7	CO916508	AGENCOURT
C 471	15.4	77.0	875	9	CC607943	OGUK046TH
C 472	15.4	77.0	883	8	CC349707	OGIA003TV
C 473	15.4	77.0	886	5	BX365216	EX365216
C 474	15.4	77.0	889	4	BI109559	603069360
C 475	15.4	77.0	893	5	BQ691775	AGENCOURT
C 476	15.4	77.0	894	5	BQ686301	AGENCOURT
C 477	15.4	77.0	895	5	BX355234	EX355234
C 478	15.4	77.0	899	5	BX355234	EX355234
C 479	15.4	77.0	910	4	BI523808	603051919
C 480	15.4	77.0	917	4	BI196513	602755631
C 481	15.4	77.0	917	5	BQ878375	AGENCOURT
C 482	15.4	77.0	920	5	BU527830	AGENCOURT
C 483	15.4	77.0	921	5	BQ422117	AGENCOURT
C 484	15.4	77.0	933	4	BG330519	602430195
C 485	15.4	77.0	933	6	BG420889	602451110
C 486	15.4	77.0	933	6	CD250920	AGENCOURT
C 487	15.4	77.0	937	4	BG828993	602753313
C 488	15.4	77.0	946	5	BQ877377	AGENCOURT
C 489	15.4	77.0	947	5	BQ690395	AGENCOURT
C 490	15.4	77.0	950	5	BX325348	EX325348
C 491	15.4	77.0	951	2	BQ642330	AGENCOURT
C 492	15.4	77.0	961	2	BE791780	601582145
C 493	15.4	77.0	965	1	AL560618	AL560618
C 494	15.4	77.0	965	4	BI199260	602758561
C 495	15.4	77.0	967	9	CC607948	OGUK046TV
C 496	15.4	77.0	968	2	BE735466	BE735466
C 497	15.4	77.0	968	9	CR169736	Forward s
C 498	15.4	77.0	971	7	CN972814	20311_124
C 499	15.4	77.0	976	4	BG341939	BG341939
C 500	15.4	77.0	984	5	BX377152	EX377152

ALIGNMENTS

RESULT 1	CB266804/c	232 bp	mRNA	linear	EST 20-FEB-2003
CH266804/c	1005710 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens				
LOCUS	CDNA 5', mRNA sequence.				
DEFINITION	CDNA 5', mRNA sequence.				
ACCESSION	CB266804.1	GI:28441390			
VERSION	EST.				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 232)				
AUTHORS	Yang, R.-Z., Shuldiner, A. and Gong, D.-W.				
TITLE	EST analysis of human adipose gene expression				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Gong Da-Wei Division of Endocrinology, Diabetes and Nutrition University of Maryland 660 Redwood St, HH497, Baltimore, MD 21201, USA Tel: 410 706 1672 Fax: 410 706 1622 Email: dgong@medicine.umaryland.edu PCR PRIMERS FORWARD: CTCGGAGCGCGCAATTCGTGTGGT BACKWARD: AATACGACTCATATAGCGGAATTGG Seq primer: GTTGGTACCGGAATTC. Location/Qualifiers 1..232 /organism="Homo sapiens"				



AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 2, 2003 this sequence version replaced gi:30340686.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqrefgenoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 9502.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0BAG028ZH06\_CS02640\_1&c=9502.r

FEATURES  
 source  
 Location/Qualifiers  
 1..610  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI027YJ02"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 100.0%; Score 20; DB 5; Length 610;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||||  
 Db 441 AGGGCGTCTCTGAGTAGCAG 422

RESULT 5  
 AY421570/c  
 LOCUS 739 bp DNA linear GSS 17-DEC-2003  
 DEFINITION Pan troglodytes HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY421570  
 VERSION AY421570.1 GI:39748429  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 1 (bases 1 to 739)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 739)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..739

gene  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 <1..>739  
 /gene="HAS1"  
 /locus\_tag="HCM7601"

ORIGIN  
 Query Match 100.0%; Score 20; DB 9; Length 739;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||||  
 Db 354 AGGGCGTCTCTGAGTAGCAG 335  
 |||||||

RESULT 6  
 BI753116/c  
 LOCUS 782 bp mRNA linear EST 25-SEP-2001  
 DEFINITION 603025961F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5196399 5',  
 mRNA sequence.  
 ACCESSION BI753116  
 VERSION BI753116.1 GI:15744694  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 782)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11491 row: j column: 16  
 High quality sequence stop: 780.  
 Location/Qualifiers  
 1..782  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5196399"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_114"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:  
 this is a NIH\_MGC Library."

FEATURES  
 source  
 Location/Qualifiers  
 1..782  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5196399"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_114"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:  
 this is a NIH\_MGC Library."

ORIGIN  
 Query Match 100.0%; Score 20; DB 4; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||||  
 Db 433 AGGGCGTCTCTGAGTAGCAG 414  
 |||||||

RESULT 7

BX371636  
 LOCUS BX371636 Homo sapiens PLACENTA COT 25-NORMALIZED EST 27-APR-2004  
 DEFINITION clone CSOD1027YJ02 3-PRIME, mRNA sequence.  
 ACCESSION BX371636  
 VERSION BX371636  
 KEYWORDS BX371636.2 GI:46616231  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 896)  
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 8, 2003 this sequence version replaced gi:30450029.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 9502.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?e=CS0BA1019ZC12\_CS01819\_1&c=9502.r

FEATURES  
 source  
 Location/Qualifiers  
 1..896  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSOD1027YJ02"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 100.0%; Score 20; DB 5; Length 896;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AGGGCGTCTCTGAGTAGCAG 20  
 Db 134 AGGGCGTCTCTGAGTAGCAG 153

RESULT 8  
 BM544718/c  
 LOCUS BM544718 1084 bp mRNA linear EST 20-FEB-2002  
 DEFINITION AGSCNCOURT 6494603 NIH\_MGC\_124 Homo sapiens cDNA clone IMAGE:5727446  
 5', mRNA sequence.  
 ACCESSION BM544718  
 VERSION BM544718.1 GI:18776197  
 KEYWORDS EST.  
 SOURCE BM544718.1 GI:18776197  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1084)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Invitrogen  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12721 row: i column: 15  
 High quality sequence start: 5  
 High quality sequence stop: 689.  
 Location/Qualifiers  
 1..1084  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5727446"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 124"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV  
 (destroyed); Site 2: NotI; RNA source male hippocampus,  
 age 27. Library is oligo-dT primed and directionally  
 cloned (EcoRV site is destroyed upon cloning). Average  
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 012."

ORIGIN  
 Query Match 100.0%; Score 20; DB 4; Length 1084;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 AGGGCGTCTCTGAGTAGCAG 20  
 Db 498 AGGGCGTCTCTGAGTAGCAG 479

RESULT 9  
 AY421569/c  
 LOCUS AY421569 1728 bp DNA linear GSS 17-DEC-2003  
 DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY421569  
 VERSION AY421569.1 GI:39748428  
 KEYWORDS GSS.  
 SOURCE AY421569.1 GI:39748428  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1728)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 1728)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 Location/Qualifiers  
 1..1728  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>1728  
 /gene="HAS1"

```

ORIGIN
Query Match          100.0%; Score 20; DB 9; Length 1728;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 1090 AGGGCGTCTCTGAGTAGCAG 1071

RESULT 10
CR602106/c
LOCUS CR602106 2072 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1027YJ02 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR602106.1 GI:50482913
VERSION CR602106
KEYWORDS HTC; CNSLT; cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 2072)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source
Query Match          90.0%; Score 18; DB 1; Length 1087;
Best Local Similarity 90.0%; Pred. No. 4.3e+02;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 927 AGGGCGTCTCTGAGTAGCAG 946

RESULT 12
BH235764/c
LOCUS BH235764 385 bp DNA linear GSS 13-NOV-2001
DEFINITION ATZKC34TF ATZK Arabidopsis thaliana genomic clone ATZKC34, genomic
survey sequence.
ACCESSION BH235764
VERSION BH235764.1 GI:16906122
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1 (bases 1 to 385)
AUTHORS Kaul,S., Town,C.D., Bowman,C.L., Van Aken,S.E., Utterback,T.V.,
Feldblyum,T.V. and Fraser,C.M.
TITLE Survey sequencing of Arabidopsis thaliana BAC T18H23
JOURNAL Unpublished (2001)
COMMENT Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
From Wash. U contig 1149. Caution: the DNA in this BAC may be from
some non-Arabidopsis source
Seq primer: TF
Class: sheared ends.
location/Qualifiers
1..385
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"

ORIGIN
Query Match          100.0%; Score 20; DB 3; Length 2072;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 1147 AGGGCGTCTCTGAGTAGCAG 1128

RESULT 11
AL571167
LOCUS AL571167 1087 bp mRNA linear EST 05-APR-2004
DEFINITION AL571167 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1027YJ02 3-PRIME, mRNA sequence.
ACCESSION AL571167
VERSION AL571167.3 GI:46237267
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1087)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31292568.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0D1027DE01NP1&c=9502.r.
location/Qualifiers
1..1087
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match          90.0%; Score 18; DB 1; Length 1087;
Best Local Similarity 90.0%; Pred. No. 4.3e+02;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 927 AGGGCGTCTCTGAGTAGCAG 946

RESULT 12
BH235764/c
LOCUS BH235764 385 bp DNA linear GSS 13-NOV-2001
DEFINITION ATZKC34TF ATZK Arabidopsis thaliana genomic clone ATZKC34, genomic
survey sequence.
ACCESSION BH235764
VERSION BH235764.1 GI:16906122
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1 (bases 1 to 385)
AUTHORS Kaul,S., Town,C.D., Bowman,C.L., Van Aken,S.E., Utterback,T.V.,
Feldblyum,T.V. and Fraser,C.M.
TITLE Survey sequencing of Arabidopsis thaliana BAC T18H23
JOURNAL Unpublished (2001)
COMMENT Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
From Wash. U contig 1149. Caution: the DNA in this BAC may be from
some non-Arabidopsis source
Seq primer: TF
Class: sheared ends.
location/Qualifiers
1..385
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"

```

/db\_xref="taxon:3702"  
/clone="ATZKC34"  
/clone\_lib="ATZK"  
/note="Vector: PH0S2; Site 1: BatX1; 2-3 kb sheared BAC  
DNA inserted into PH0S2 using BatX1 linkers"

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 385;  
Best Local Similarity 94.7%; Pred. No. 8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
|||||  
Db 101 GGGCTTCTCTGAGTAGCAG 83

RESULT 13  
AZ408342/c

LOCUS AZ408342 521 bp DNA linear GSS 03-OCT-2000

DEFINITION IM0179D01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0179D01 R, genomic survey sequence.

ACCESSION AZ408342

VERSION AZ408342.1 GI:10532355

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 521)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Becorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0179 row: D column: 01  
Seq primer: CACAGGGAACGATGACC  
Class: plasmid ends  
High quality sequence stop: 521.  
Location/Qualifiers  
1..521  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0179D01"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnases/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 87.0%; Score 17.4; DB 8; Length 521;  
Best Local Similarity 94.7%; Pred. No. 8.2e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCA 19  
|||||  
Db 76 AGGGCGTCTCTGAGTAGCA 58

RESULT 14  
CO708502

LOCUS CO708502 578 bp mRNA linear EST 27-JUL-2004

DEFINITION DG14-10b23 DG14-muscle Canis familiaris cDNA 3', mRNA sequence.

ACCESSION CO708502

VERSION CO708502.1 GI:50673643

KEYWORDS EST.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 578)

AUTHORS Schluter, T., Hermanns, J., Weindel, M., Schuette, D., Kranz, H., Henrich, J. and Loebbert, R.

TITLE Dog arrayTAG cDNA clone collection

JOURNAL Unpublished (2004)

COMMENT Contact: Thomas Schluter  
LION Bioscience AG  
Walhoferstrasse 98, D-69123 Heidelberg, Germany  
Tel: +49 6221 4038 150  
Fax: +49 6221 4038 290  
Email: Thomas.Schluter@lionbioscience.com.  
Location/Qualifiers  
1..578  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/strain="Beagle"  
/db\_xref="taxon:9615"  
/tissue\_type="muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="DG14-muscle"  
/note="Organ: muscle; Vector: Dog pBluescript LION"

FEATURES  
source

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 578;  
Best Local Similarity 94.7%; Pred. No. 8.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
|||||  
Db 487 GGGCGTCTCTGAGTAGGAG 505

RESULT 15  
BH236262/c

LOCUS BH236262 593 bp DNA linear GSS 13-NOV-2001

DEFINITION ATZKE45TF ATZK Arabidopsis thaliana genomic clone ATZKE45, genomic survey sequence.

ACCESSION BH236262

VERSION BH236262.1 GI:16906620

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (chale cress)

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.





Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 210)

REFERENCE  
AUTHORS  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL  
COMMENT  
Unpublished (2000)

CONTACT: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00  
Plate: 0152 row: P column: 11  
Seq primer: CTTGTAAACGACGCCCAT  
Class: plasmid ends  
High quality sequence stop: 210.

FEATURES  
source  
Location/Qualifiers  
1..210  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0152P11"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN  
Query Match 84.0%; Score 16.8; DB 8; Length 210;  
Best Local Similarity 90.0%; Pred. No. 1.5e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 82 AGGGCGTCTCTGAGTAGCAG 101

RESULT 19  
CR037148  
LOCUS CR037148 249 bp DNA linear GSS 05-JUL-2004  
DEFINITION Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN307118, genomic survey sequence.

ACCESSION CR037148  
VERSION CR037148.1 GI:4970203  
KEYWORDS GSS; genome survey sequence; MICER.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 249)

REFERENCE  
AUTHORS  
Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L., Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y., Rogers, J. and Bradley, A.

TITLE  
Direct Submission

JOURNAL  
COMMENT  
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES  
source  
Location/Qualifiers  
1..249  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN307118"  
/clone\_lib="MHPN"

ORIGIN  
Query Match 84.0%; Score 16.8; DB 9; Length 249;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 187 AGGGAGTCTCTGAGTAGCAG 168

RESULT 20  
CE277790/c  
LOCUS CE277790 252 bp DNA linear GSS 26-SEP-2003  
DEFINITION tigr-gss-dog-17000333680929 Dog Library Canis familiaris genomic, genomic survey sequence.

ACCESSION CE277790  
VERSION CE277790.1 GI:36025605  
KEYWORDS GSS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE  
AUTHORS  
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and Venter, J.C.

TITLE  
The dog genome: survey sequencing and comparative analysis

JOURNAL  
MEDLINE  
PUBMED  
22875432  
14512627

COMMENT  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org  
Class: shotgun.

FEATURES  
source  
Location/Qualifiers  
1..252  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BetXI; Libraries were prepared from peripheral blood"

ORIGIN  
Query Match 84.0%; Score 16.8; DB 9; Length 252;  
Best Local Similarity 90.0%; Pred. No. 1.6e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 56 AGGGCGTCTCTGAGTAGCAG 37

```

RESULT 21
CR012636/c
LOCUS
DEFINITION
  CR012636 402 bp DNA linear GSS 05-JUL-2004
  Reverse strand read from insert in 3'HPRT insertion targeting and
  chromosome engineering clone MHP341k19, genomic survey sequence.
ACCESSION
  CR012636.1 GI:49745683
VERSION
  GSS; genome survey sequence; MICE.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
  1 (bases 1 to 402)
AUTHORS
  Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
  Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
  Rogers,J. and Bradley,A.
DIRECT SUBMISSION
  Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. http://www.sanger.ac.uk/MICER
JOURNAL
  Location/Qualifiers
FEATURES
  source
  1..402
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /clone="MHP341k19"
  /clone_lib="MHPP"

ORIGIN
  Query Match 84.0%; Score 16.8; DB 9; Length 402;
  Best Local Similarity 90.0%; Pred. No. 1.6e+03;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
  ||||| ||||| ||||| |||||
Db 317 AGGGATCTCTGAGTAGCAG 298

RESULT 22
R23689/c
LOCUS
DEFINITION
  YQ32h10.r1 Soares infant brain 1N1B Homo sapiens cDNA clone
  IMAGE:34208 5', mRNA sequence.
ACCESSION
  R23689.1 GI:778577
VERSION
  EST.
KEYWORDS
  Homo sapiens (human)
SOURCE
  Homo sapiens
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 464)
REFERENCE
  Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
  Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
  Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,P.,
  Trevaskis,E., Waterson,R., Williamson,A., Wohldmann,P. and
  Wilson,R.
  The WashU-Merck EST Project
  Contact: Wilton RK
  Unpublished (1995)
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@watson.wustl.edu
  Insert Size: 1583
  High quality sequence stops: 298 Source: IMAGE Consortium, LLNL
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1583 Std Error: 0.00
  Seq primer: M13RP1
  High quality sequence stop: 298.
  Location/Qualifiers
  1..464

FEATURES
  source
  1..464

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:406555"
/db_xref="taxon:9606"
/clone="IMAGE:34208"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain 1N1B"
/clone="Organ: whole brain; Vector: Lafmid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(GT) primer [5',
AACTGGAAGATTCGCGCCGACGAGATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lafmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN
  Query Match 84.0%; Score 16.8; DB 7; Length 464;
  Best Local Similarity 90.0%; Pred. No. 1.6e+03;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
  ||||| ||||| ||||| |||||
Db 67 AGGGAGTCTCTGAGAGCAG 48

RESULT 23
A0229055
LOCUS
DEFINITION
  HS_2019_A1_H06_MR CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=2019 Col=11 Row=O, genomic survey
  sequence.
ACCESSION
  A0229055
VERSION
  A0229055.1 GI:3654284
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 519)
REFERENCE
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  99380589
  10449784
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Sequence Tagged Connector
  Plate: 2019 row: O column: 11
  Class: BAC ends
  High quality sequence stop: 519.
  Location/Qualifiers
  1..519
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone="Plate=2019 Col=11 Row=O"
  /sex="male"
  /clone_lib="CIT Approved Human Genomic Sperm Library D"
  /note="Organ: sperm; Vector: pBelosAC11; BAC Clones in
  E-Coli DH10B"

ORIGIN

```

```

Query Match      84.0%; Score 16.8; DB 8; Length 519;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 283 AGGCATCTCTGAGTTGCAG 302

RESULT 24
CR265831/c
LOCUS
DEFINITION
Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN303b23, genomic survey sequence.
CR265831
ACCESSION
CR265831.1 GI:50044684
VERSION
GSS; genome survey sequence; MICER.
KEYWORDS
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 547)
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
TITLE
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgehire,
JOURNAL
CB10 18A, UK. http://www.sanger.ac.uk/MICER
FEATURES
source
Location/Qualifiers
1..547
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN303b23"
/clone_lib="MHPN"

ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 547;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 540 AGGAGTCACTGAGTAGCAG 521

RESULT 25
CE015183/c
LOCUS
DEFINITION
tigr-gss-dog-17000321305524 Dog Library Canis familiaris genomic,
genomic survey sequence.
CE015183
ACCESSION
CE015183.1 GI:35021658
VERSION
GSS.
KEYWORDS
Canis familiaris (dog)
SOURCE
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 550)
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
PUBMED
CONTACT
Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208

Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1..550
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strains="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

FEATURES
source
Location/Qualifiers
1..551
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="aof02-9ms4-f04"
/tissue type="female inflorescences"
/lab host="SOLR"
/clone_lib="Aof02"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at http://fgp.bio.psu.edu"

ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 550;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 217 AGGAGTCTCAGAGTAGCAG 198

RESULT 26
CV458871/c
LOCUS
DEFINITION
aof02-9ms4-f04 Aof02 Asparagus officinalis cDNA clone
aof02-9ms4-f04 5', mRNA sequence.
CV458871
ACCESSION
CV458871.1 GI:52846057
VERSION
EST.
KEYWORDS
Asparagus officinalis (garden asparagus)
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.
1 (bases 1 to 551)
dePamphilis,C., Carlson,J., Ma,H., Soltis,D., Soltis,P.,
Oppenheimer,D., Frohlich,M., Doyle,J., Tankalev,S., Webb,M.,
Leebens-Mack,J., Landherr,L., Ilut,D. and Wall,K.
Generation of ESTs from early female inflorescences of Asparagus
officinalis
Unpublished (2004)
Contact: Claude dePamphilis or James Leebens-Mack
Mueller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl110@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (http://pgn.cornell.edu)
Plate: aof02-9ms4 row: f column: 04
Seq primer: M13F.

FEATURES
source
Location/Qualifiers
1..551
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="aof02-9ms4-f04"
/tissue type="female inflorescences"
/lab host="SOLR"
/clone_lib="Aof02"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at http://fgp.bio.psu.edu"

ORIGIN

```

```

Query Match      84.0%; Score 16.8; DB 7; Length 551;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 161 AGGGCGTCTCAGAGTGCGAG 142

RESULT 27
LOCUS BI788920
DEFINITION ie41c05.y1 Kaestner ngn3 wt Mus musculus cDNA 5', mRNA linear EST 01-OCT-2001
ACCESSION BI788920
VERSION BI788920.1 GI:15816645
SOURCE EST.
ORGANISM Mus musculus (house mouse)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
COMMENT 1 (bases 1 to 554)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Fillier, L., Maiza, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,
Williams, T., Jackson, F. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ie41c05.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
2000) Library was constructed by Catherine Lee DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Marie Searce
(msearce@mail.med.upenn.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 434.
FEATURES
source
    Location/Qualifiers
        1..554
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="129/Sv x CD1"
            /db_xref="taxon:10090"
            /dev_stage="p.c. 14.5"
            /lab_host="E. coli-DH12S (GIBCO)"
            /clone_lib="Kaestner ngn3 wt"
            /note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1:
            Not 1; Site_2: Sal I; The library was prepared by
            Catherine S. Lee and has not been published. The pancreas
            was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
            2000). The cDNA's were prepared with an oligo containing a
            NotI site, and SalI linkers were added to the ends. The
            inserts were cut with NotI before being cloned into the
            NotI-SalI sites in the vectors. This is one of two
            libraries, ngn3 wt and ngn3 -/- . The wt library is in
            pSPORT1, T7 promoter is 5'."

Query Match      84.0%; Score 16.8; DB 4; Length 554;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 70 AGGGTTCTCTGAGTAGCAG 89

RESULT 28
LOCUS CA248736
DEFINITION SCSBFL1104B09.g FL1 Saccharum officinarum cDNA clone SCSBFL1104B09
ACCESSION CA248736
VERSION CA248736.1 GI:35330676
SOURCE EST.
ORGANISM Saccharum officinarum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Saccharum officinarum
COMMENT The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br
Plate: 104 row: B column: 09
Seq primer: T7 Promoter Primer.
FEATURES
source
    Location/Qualifiers
        1..565
            /organism="Saccharum officinarum"
            /mol_type="mRNA"
            /db_xref="taxon:4547"
            /clone="SCSBFL1104B09"
            /lab_host="DH10B"
            /clone_lib="FL1"
            /note="Organ: Inflorescence at beginning of development
            (1cm-long); Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
            An unidirectional cDNA library generated from
            Inflorescence at beginning of development (1cm-long)).
            cDNA was prepared from polyA+ mRNA using Superscript
            Plasmid System Kit (Invitrogen). The double-strand cDNAs
            were fractionated in a sepharose CL-2B 40cm-columns and
            fragments sizing between 0.8 and 1.5 Kb were
            directionally cloned into the vector. Details of each
            source of RNA and library construction can be obtained at
            http://sucst.lad.ic.unicamp.br/public"

Query Match      84.0%; Score 16.8; DB 6; Length 565;
Best Local Similarity 90.0%; Pred. No. 1.6e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 10 AGGGCGTCTCGGAGGAGCAG 29

RESULT 29
LOCUS BM335950/c
DEFINITION MEST179-D02.T3 ISUM5-RN Zea mays cDNA clone MEST179-D02 3', mRNA
ACCESSION BM335950
VERSION BM335950.1 GI:18166111
SOURCE EST.
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 569)

Wen.T.J., Qiu.P., Guo.L., Ashlock.D.A and Schnable.P.S.

Expressed Sequence Tags from B73 Maize: various stages and tissues

Including seedlings treated with a variety of hormones

Unpublished (2001)

Contact: Patrick S. Schnable

Schnable Laboratory

Iowa State University

G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA

Tel: 515-294-0975

Fax: 515-294-2299

Email: schnable@iastate.edu

Individual basecall and confidence value were assigned using the

Phred software.

(<http://depts.washington.edu/ventures/collabtr/direct/index.htm#b>

rt). Overall sequence quality assessment and vector trimming were

conducted using the Lucy software (<http://www.tigr.org/softlab/lucy>).

Lucy parameters were set to ensure an overall trimmed quality of

97.5% or better without any vector fragments in the chosen

high-quality region of each sequence. Low-quality bases between the

poly-T and the high-quality region were replaced with N's to serve

as spacers.

PCR primers

FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)

BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

Location/Qualifiers

#### FEATURES

source

1..569

/organism="Zea mays"

/mol\_type="mRNA"

/cultivar="B73"

/db\_xref="taxon:4577"

/clone="MEST179-D02"

/tissue\_type="mixed"

/lab\_host="DH10B"

/clone\_lib="ISUM5-RN"

/note="Vector: pT7T3PAC; Site 1: EcoRI; Site 2: NotI;

Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),

Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels

(3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65

DAG), Tassel (3-39 cm, 53 and 56 DAG), Immature ear

(0.2-3.0 cm, 53, 56, 59 DAG), Husk (73 DAG), Silk,

unpollinated first ear, ear shank, etiolated seedlings,

callus, Cycloheximide-treated callus, Anaerobic treated

seedlings, NAA (a-Naphthalene acetic acid)-treated

seedlings, Kinetin-treated seedlings, ACPG

(1-aminocyclopropane-1-carboxylic acid)-treated seedlings,

Brassinolide-treated seedlings, ABA (Abscisisic

acid)-treated seedlings, GA (Gibberellic acid)-treated

seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA

molecules were generated as follows. First-strand cDNA was

prepared from oligo-dT selected mRNA by priming with a

NotI oligo-dT primer (5'

NACTGGAAGATTCGGCGCGAGCAATTTTTTTTTTTTTTTT). The

resulting DNA:RNA hybrid was treated with RNase H and used

as a template for DNA PolI-catalyzed second strand

synthesis. After the addition of EcoRI adaptors, the

ds-cDNAs were digested with NotI and size-selected. The

resulting molecules were directionally cloned into the

EcoRI and NotI sites of the pT7T3PAC vector. The library

then went through one round of normalization to Cot value

of 5 based on the methods of Marcelo Bento Soares (Genome

Research 6: 791-806, 1996)."

#### ORIGIN

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 4; Length 569;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20

||||| |||||||

Db 186 ATGGCGTTTCTGAGTAGCAG 167

#### RESULT 30

CG985592/c

LOCUS

DEFINITION

CG985592

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

CG985592 593 bp DNA linear GSS 15-DEC-2003  
CH240\_155N13.TV CHORI-240 Bos taurus genomic clone CH240\_155N13,  
genomic survey sequence.

CG985592

CG985592.1 GI:39911371

GSS.

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

CG985592 593 bp DNA linear GSS 15-DEC-2003  
CH240\_155N13.TV CHORI-240 Bos taurus genomic clone CH240\_155N13,  
genomic survey sequence.

CG985592

CG985592.1 GI:39911371

GSS.

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

DEFINITION BOHZN67TF BO 2.3 KB Brassica oleracea genomic clone BOHZN67, genomic survey sequence.

ACCESSION BH646059  
VERSION BH646059.1 GI:18703452  
KEYWORDS GSS.

SOURCE Brassica oleracea  
ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 645)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

FEATURES

source

1. .645

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOHZN67"

/clone\_lib="BO 2.3 KB"

/notes="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHO51 using BstXI linkers"

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 645;  
Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20

||||| |||||||||

DB 354 AGGCTCTCTGAGTAGCAG 335

RESULT 32

AZ966847/C

LOCUS

DEFINITION 2M0237106R Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0237106 R, genomic survey sequence.

ACCESSION AZ966847

VERSION AZ966847.1 GI:13838074

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 647)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0237 row: I column: 06

Seq primer: CACACAGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 647.

Location/Qualifiers

1. .647

source

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0237106"

/sex="Female"

/lab\_host="E. coli strain XL10-Gold, TI-resistant, F-"

/clone\_lib="Mouse 10kb plasmid UUGC2M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (female) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 647;

Best Local Similarity 90.0%; Pred. No. 1.7e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20

||||| |||||||||

DB 164 AGGCTTCTCTGAGTAGCAG 145

RESULT 33

AZ840734/C

LOCUS

DEFINITION 2M0138122F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0138122 F, genomic survey sequence.

ACCESSION AZ840734

VERSION AZ840734.1 GI:13010642

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 664)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0138 row: I column: 22  
 Seq primer: CCGTTGTAACACGCGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 664.

#### FEATURES

Location/Qualifiers  
 1..664  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUGC2M0138122"  
 /sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Ti-resistant, F-"  
 /clone\_lib="Mouse 10kb plasmid UUGC1M library"  
 /note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 664;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 ||||| ||||| ||||| |||||  
 Db 181 AGGGCGACTCGGAGTAGCAG 162

#### RESULT 34

CV288705  
 LOCUS 669 bp mRNA linear EST 23-SEP-2004  
 DEFINITION aof01-9ms3-a04 Aof01 Asparagus officinalis cDNA clone  
 aof01-9ms3-a04 5', mRNA sequence.

CV288705  
 CV288705.1 GI:52572310

EST.

Asparagus officinalis (garden asparagus)

#### SOURCE

Asparagus officinalis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;  
 Asparagaceae; Asparagus.

1 (bases 1 to 669)

dePamphilis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,  
 Oppenheimer, D., Frohlich, M., Doyle, J., Tankley, S., Webb, M.,  
 Leebens-Mack, J., Landherr, L., Ilut, D. and Wall, K.

Generation of ESTs from early male inflorescences of Asparagus

officinalis

Unpublished (2004)

Contact: Claude dePamphilis or James Leebens-Mack

Mueller Laboratory

Penn State University

208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn

State University, University Park, PA 16802, USA

Tel: 814 863 6413

Fax: 814 865 9131

Email: cwd3@psu.edu or jhl10@psu.edu

The sequence provided is trimmed of vector and low quality regions.

Full sequence and original trace file are available from the Plant  
 Genome Network website (http://pgn.cornell.edu)  
 Plate: aof01-9ms3 row: a Column: 04  
 Seq primer: M13F.

#### FEATURES

Location/Qualifiers  
 1..669  
 /organism="Asparagus officinalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4686"  
 /clone="aof01-9ms3-a04"  
 /tissue\_type="male inflorescences"  
 /lab\_host="SOLR"  
 /clone\_lib="Aof01"

/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;  
 Site 2: XhoI; This is a directionally cloned,  
 non-normalized library. This library has been generated by  
 the Floral Genome Project (FGP). The Floral Genome Project  
 is funded by NSF's Plant Genome Research Program  
 (DBI-0115684). More information about the project can be  
 obtained at http://fgp.bio.psu.edu"

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 669;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 ||||| ||||| ||||| |||||  
 Db 380 AGGGCGTCTCAGATGGCAG 399

#### RESULT 35

CR191252/c

LOCUS 679 bp DNA linear GSS 06-JUL-2004

DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and  
 chromosome engineering clone MHPN1109, genomic survey sequence.

ACCESSION CR191252

VERSION CR191252.1 GI:49970101

KEYWORDS GSS; genome survey sequence; MICER.

SOURCE Mus musculus (house mouse)

#### ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 679)

Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,

Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,

Rogers, J. and Bradley, A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

#### FEATURES

Location/Qualifiers  
 1..679  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPN1109"  
 /clone\_lib="MHPN"

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 679;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 ||||| ||||| ||||| |||||  
 Db 571 AGGGAGTCACTGAGTAGCAG 552

#### RESULT 36

AV370109

LOCUS 684 bp mRNA linear EST 24-OCT-2001

DEFINITION AV370109 RIKEN full-length enriched, adult male colon Mus musculus

cDNA clone 9030207M06 3', mRNA sequence.

```

ACCESSION AV370109
VERSION AV370109.2 GI:16397792
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 684)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Nov 14, 1999 this sequence version replaced gi:6417756.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yananaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
FEATURES
source
1. .684
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9030207M06"
/sex="male"
/tissue_type="colon"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male colon"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCTCGAGTTAATTAATCCCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end:
BamHI"
ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 684;
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGATGACGAG 20
Db 331 AGGGTTCTCTGAGTAGCAG 350
RESULT 37
AG119191
LOCUS AG119191 702 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-127124.F, genomic survey sequence.
ACCESSION AG119191
VERSION AG119191.1 GI:16739710
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 702)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbescgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .702
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-127124.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 84.0%; Score 16.8; DB 9; Length 702;
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AGGGCGTCTCTCGATGACGAG 20

```





The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@igr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library  
 availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org)).  
 Clones may be purchased from BACPAC Resources  
 ([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end  
 page: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)  
 Plate: 40 row: G column: 3  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES

source Location/Qualifiers  
 1..730  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SsNHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-40G3"  
 /sex="female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 1"  
 /note="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;  
 CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
 Pieter de Jong"

# ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 730;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 114 AGGGAGTGTCTGAGTAGCAG 133

RESULT 41  
 AG383658/c  
 LOCUS 736 bp DNA linear GSS 03-JUN-2004  
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-194019.T7, genomic survey  
 sequence.  
 ACCESSION AG383658  
 VERSION AG383658.1 GI:47994863  
 KEYWORDS GSS.  
 SOURCE Mus musculus molossinus  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 BAC end Sequences of Library MSMg01

REFERENCE  
 AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 [E-mail:hattori@gsr.riken.jp, URL:<http://hgp.gsc.riken.go.jp/>,  
 Tel:81-45-503-9111, Fax:81-45-503-9170]  
 Clones are derived from the mouse BAC library MSMg01. For BAC  
 library availability, please contact Kuniya Abe ([abe@rtc.riken.jp](mailto:abe@rtc.riken.jp)).  
 Tsukuba Institute, Bio Resource Center,  
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Koyadai, Tsukuba, 305-0074 Japan  
 phone: 81-298-36-9189, fax: 81-298-36-9199  
 e-mail: abe@rtc.riken.jp  
 PRIMERS  
 Sequencing : T7  
 LIBRARY Vector : pBACe3.6

# COMMENT

R.Site 1 : EcoRI  
 R.Site 2 : EcoRI.  
 Location/Qualifiers  
 1..736  
 /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSMg01-194019.T7"  
 /sex="male"  
 /tissue\_type="mixture of kidney and spleen"  
 /clone\_lib="MSMg01 Mouse Male BAC Library"

# ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 736;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGGCTCTCTGAGTAGCAG 20  
 |||||  
 Db 82 AGGGGTCTCTGAGTAGCTG 63

RESULT 42  
 AG569158/c

LOCUS 744 bp DNA linear GSS 05-JUN-2004  
 DEFINITION Mus musculus molossinus DNA, clone:MSMg01-491B06.TJ, genomic survey  
 sequence.  
 ACCESSION AG569158  
 VERSION AG569158.1 GI:48329856  
 KEYWORDS GSS.  
 SOURCE Mus musculus molossinus  
 ORGANISM Mus musculus molossinus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 BAC end Sequences of Library MSMg01

REFERENCE  
 AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 [E-mail:hattori@gsr.riken.jp, URL:<http://hgp.gsc.riken.go.jp/>,  
 Tel:81-45-503-9111, Fax:81-45-503-9170]  
 Clones are derived from the mouse BAC library MSMg01. For BAC  
 library availability, please contact Kuniya Abe ([abe@rtc.riken.jp](mailto:abe@rtc.riken.jp)).  
 Tsukuba Institute, Bio Resource Center,  
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
 Koyadai, Tsukuba, 305-0074 Japan  
 phone: 81-298-36-9189, fax: 81-298-36-9199  
 e-mail: abe@rtc.riken.jp  
 PRIMERS  
 Sequencing : TJ  
 LIBRARY Vector : pBACe3.6  
 R.Site 1 : EcoRI.  
 R.Site 2 : EcoRI.

# FEATURES

source Location/Qualifiers  
 1..744  
 /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSMg01-491B06.TJ"  
 /sex="male"  
 /tissue\_type="mixture of kidney and spleen"  
 /clone\_lib="MSMg01 Mouse Male BAC Library"

# ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 744;

```

Best Local Similarity 90.0%; Pred. No. 1.7e+03; Mismatches 2; Indels 0; Gaps 0;
Matches 18; Conservative 0;

QY 1 AGGGGCTCTCTGAGTAGCAG 20
||||| ||||||| |||
Db 290 AGGGCTCTCTGAGTAACAG 271

RESULT 43
CC552929
LOCUS CH240_438B16.TARBAC13P2 CHORI-240 Bos taurus genomic clone
DEFINITION CH240_438B16, genomic survey sequence.
ACCESSION CC552929
VERSION CC552929.1 GI:318711213
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 744)
AUTHORS Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M., Chiu,R.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: CH240_438B16.T7
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bccgc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering.information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 438 row: B column: 16
Seq primer: SP6
Class: BAC ends.
FEATURES             Location/Qualifiers
     source            1..744
     organism="Bos taurus"
     /mol_type="genomic DNA"
     /strain="breed: Hereford"
     /db_xref="taxon:9913"
     /clone="CH240_438B16"
     /sex="Male"
     /cell_type="Blood"
     /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
     Hereford bull I1 Domino 99375; CHORI-240 Bovine BAC
     library (Male) produced by Pieter de Jong"

ORIGIN
Query Match      84.0%; Score 16.8; DB 9; Length 744;
Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGGCTCTCTGAGTAGCAG 20
||||| ||||||| |||
Db 460 AGGGTTCTCTGAGTAGCAG 441

RESULT 45
AG591654
LOCUS AG591654
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-520K21.T7, genomic survey
sequence.
ACCESSION AG591654
VERSION AG591654.1 GI:48352484
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 773)
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Sushiro-cho,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.

```

```

FEATURES
  source
    Location/Qualifiers
      1..773
        /organism="Mus musculus molossinus"
        /mol_type="genomic DNA"
        /sub_species="molossinus"
        /db_xref="taxon:57486"
        /clone="MSMg01-520K21.T7"
        /sex="male"
        /tissue_type="mixture of kidney and spleen"
        /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
  Query Match
  Best Local Similarity 84.0%; Score 16.8; DB 9; Length 773;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 605 AGGGCGTCTCTGAGGAG 624

RESULT 46
CV292304
LOCUS
DEFINITION
  CV292304 774 bp mRNA linear EST 23-SEP-2004
aof01-11ms3-h11 Aof01 Asparagus officinalis cDNA clone
CV292304
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Asparagus officinalis (garden asparagus)
  Asparagus officinalis
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
  Asparagaceae; Asparagus.
  1 (bases 1 to 774)
  Oppenheimer,D., Frohlich,M., Doyle,J., Tankesley,S., Webb,M.,
  Leebens-Mack,J., Landherr,L., Ilut,D. and Wall,K.
  Generation of ESTs from early male inflorescences of Asparagus
  officinalis
  Unpublished (2004)
  Contact: Claude dePamphilis or James Leebens-Mack
  Mueller Laboratory
  Penn State University
  208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
  State University, University Park, PA 16802, USA
  Tel: 814 863 6413
  Fax: 814 865 9131
  Email: cwd3@psu.edu or jhl10@psu.edu
  The sequence provided is trimmed of vector and low quality regions.
  Full sequence and original trace file are available from the Plant
  Genome Network website (http://pgn.cornell.edu)
  Plate: aof01-11ms3 row: h column: 11
  Seq primer: M13F.

FEATURES
  source
    Location/Qualifiers
      1..774
        /organism="Asparagus officinalis"
        /mol_type="mRNA"
        /db_xref="taxon:4686"
        /clone="aof01-11ms3-h11"
        /tissue_type="male inflorescences"
        /lab_host="SOLR"
        /clone_lib="Aof01"
        /notes="Vector: pBluescript SK (+/-); Site 1: EcoRI;
        Site 2: XhoI; This is a directionally cloned,
        non-normalized library. This library has been generated by
        the Floral Genome Project (FGP). The Floral Genome Project
        is funded by NSF's Plant Genome Research Program
        (DBI-0115684). More information about the project can be
        obtained at http://fgp.bio.psu.edu"

ORIGIN
  Query Match
    84.0%; Score 16.8; DB 7; Length 774;

```

---

```

Best Local Similarity 90.0%; Pred. No. 1.7e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 438 AGGGCGTCTCTGAGTGGCAG 457

RESULT 47
CL662911
LOCUS
DEFINITION
  CL662911 796 bp DNA linear GSS 09-JUL-2004
PRI0142d All - PRI0142d.B21 (796) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL662911
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Pristionchus pacificus
  Pristionchus pacificus
  Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
  Neodiplogasteridae; Pristionchus.
  1 (bases 1 to 796)
  Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  AppADB: an AcedB database for the nematode satellite organism
  Pristionchus pacificus
  Nucleic Acids Res. 32 (1), D421-D422 (2004)
  Contact: Sommer RJ
  Evolutionary Biology
  Max-Planck-Institute for Developmental Biology
  Spemannstr. 37-39, Tuebingen D-72076, Germany
  Tel: 00497071601371
  Fax: 00497071601498
  Email: raif.sommer@tuebingen.mpg.de
  This library was generated at Caltech, Pasadena, USA and end
  sequenced at Vancouver, Canada.
  Seq primer: T7
  Class: fosmid ends.
  Location/Qualifiers
    1..796
      /organism="Pristionchus pacificus"
      /mol_type="genomic DNA"
      /strain="California"
      /db_xref="taxon:54126"
      /clone_lib="Mixed stage fosmid library of P. pacificus
      var. California"
      /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
  Query Match
    84.0%; Score 16.8; DB 9; Length 796;
  Best Local Similarity 90.0%; Pred. No. 1.7e+03;
  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 777 AGCGCGTCTCTGAGTAGCAG 796

RESULT 48
CV289382/c
LOCUS
DEFINITION
  CV289382 846 bp mRNA linear EST 23-SEP-2004
aof01-2ms4-el2 Aof01 Asparagus officinalis cDNA clone
aof01-2ms4-el2 5', mRNA sequence.
CV289382
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Asparagus officinalis (garden asparagus)
  Asparagus officinalis
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
  Asparagaceae; Asparagus.
  1 (bases 1 to 846)
  dePamphilis,C., Frohlich,M., Doyle,J., Tankesley,S., Webb,M.,
  Oppenheimer,D.,

```

**TITLE**  
**JOURNAL**  
**COMMENT**

Leebens-Mack, J., Landherr, L., Ilut, D. and Wall, K.  
 Generation of ESTs from early male inflorescences of Asparagus officinalis  
 Unpublished (2004)  
 Contact: Claude dePamphilis or James Leebens-Mack  
 Mueller Laboratory  
 Penn State University  
 208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn State University, University Park, PA 16802, USA  
 Tel: 814 863 6413  
 Fax: 814 865 9131  
 Email: cwd3@psu.edu or jhl10@psu.edu  
 The sequence provided is trimmed of vector and low quality regions. Full sequence and original trace file are available from the Plant Genome Network website (<http://pgn.cornell.edu>)  
 Plate: aof01-2ms4 row: e column: 12  
 Seq primer: M13F.

**FEATURES**  
 source  
 1. .846  
 /organism="Asparagus officinalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4686"  
 /clone="aof01-2ms4-e12"  
 /tissue\_type="male inflorescences"  
 /lab\_host="SOLR"  
 /clone\_lib="Aof01"  
 /notes="Vector: pBluescript SK (+/-); Site 1: EcoRI; Site 2: XhoI; This is a directionally cloned, non-normalized library. This library has been generated by the Floral Genome Project (FGP). The Floral Genome Project is funded by NSF's Plant Genome Research Program (DBI-0115684). More information about the project can be obtained at <http://fgp.bio.psu.edu>"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 7; Length 846;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||  
 DB 267 AGGGCGTCTCAGAGTGGCAG 248

**RESULT 49**  
**CR802378/c**  
**LOCUS**  
**DEFINITION**  
 CR802378 855 bp DNA linear GSS 24-SEP-2004  
 GR0AAA21DB07FM1 INRA BAC Bos taurus genomic clone INRA\_425C11, DNA sequence, genomic survey sequence.

**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**

GSS.  
 Bos taurus (cow)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**

1 (bases 1 to 855)  
 Eggen, A., Schibler, L. and Roy, A.  
 Bovine BAC End Sequences from the INRA bovine BAC library  
 Unpublished  
 2 (bases 1 to 855)  
 Direct Submission  
 Submitted (20-SEP-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Contact: Andre Eggen  
 Department of Animal Genetics - LGBC  
 INRA

## COMMENT

78350 Jouy-en-Josas, France  
 Tel: 33 1 34 65 24 24  
 Fax: 33 1 34 65 24 78

Email: [eggen@jouy.inra.fr](mailto:eggen@jouy.inra.fr)  
 Clones are derived from the INRA bovine BAC library ([http://locus.jouy.inra.fr/fpc/cattle\\_bac\\_map.htm](http://locus.jouy.inra.fr/fpc/cattle_bac_map.htm)). For BAC library availability, please contact Andre Eggen ([eggen@jouy.inra.fr](mailto:eggen@jouy.inra.fr)). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by INRA (Jouy-en-Josas) and Genoscope (Evry) Plate: 425 row: C column: 11  
 Seq primer: M13 Forward  
 Class: BAC ends.

**FEATURES**  
 Location/Qualifiers  
 source  
 1. .855  
 /organism="Bos taurus"  
 /mol\_type="genomic DNA"  
 /strain="breed: Holstein"  
 /db\_xref="taxon:9913"  
 /clone="INRA\_425C11"  
 /sex="Male"  
 /cell\_type="fibroblast"  
 /clone\_lib="INRA bovine BAC"  
 /note="Vector: pBelOAC11; Site 1: HindIII; Holstein bull; INRA Bovine BAC library (Male) produced by Andre Eggen-Genoscope sequence ID : GR0AAA21DB07FM1"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 855;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||  
 DB 727 AGGGCGTCTCTGAGTGGCAG 708

## RESULT 50

**CR194037/c**  
**LOCUS**  
**DEFINITION**  
 CR194037 874 bp DNA linear GSS 06-JUL-2004  
 Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN354n23, genomic survey sequence.

**ACCESSION**  
**VERSION**  
**KEYWORDS**  
**SOURCE**  
**ORGANISM**

CR194037.1 GI:49972886  
 GSS; genome survey sequence; MICER.  
 Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**

1 (bases 1 to 874)  
 Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L., Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y., Rogers, J. and Bradley, A.  
 Direct Submission  
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>

## FEATURES

Location/Qualifiers  
 source  
 1. .874  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHPN354n23"  
 /clone\_lib="MHPN"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 874;  
 Best Local Similarity 90.0%; Pred. No. 1.7e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
 |||||  
 DB 571 AGGGAGTCACTGAGTAGCAG 552

Search completed: March 14, 2005, 21:20:08  
 Job time : 1968 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 17:14:07 ; Search time 102 Seconds  
(without alignments)  
320.839 Million cell updates/sec

Title: US-10-672-399-10

Perfect score: 20

Sequence: 1 agggcgctctctgagtagcag 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	2087	4	US-09-949-016-5659
C 2	20	100.0	2116	3	US-08-635-552A-1
C 3	20	100.0	2117	3	US-09-155-769-1
C 4	20	100.0	14862	4	US-09-949-016-17401
5	16.8	84.0	278	4	US-09-513-999C-14085
6	16.8	84.0	313	4	US-09-513-999C-2290
7	15.8	79.0	525	4	US-09-904-196B-2
8	15.8	79.0	525	4	US-09-760-008A-2
9	15.8	79.0	525	4	US-10-192-294-2
10	15.8	79.0	601	4	US-09-949-016-62651
11	15.8	79.0	601	4	US-09-949-016-62652
12	15.8	79.0	601	4	US-09-949-016-62687
13	15.8	79.0	601	4	US-09-949-016-62688
14	15.8	79.0	601	4	US-09-949-016-88843
15	15.8	79.0	601	4	US-09-949-016-88844
16	15.8	79.0	882	4	US-09-253-991A-10864
17	15.8	79.0	972	4	US-09-902-540-5386
18	15.8	79.0	1203	4	US-09-252-991A-10820
19	15.8	79.0	1315	4	US-09-902-540-215
20	15.8	79.0	2614	2	US-08-795-868-15
21	15.8	79.0	2614	3	US-09-303-069-15
22	15.8	79.0	2614	3	US-09-134-250-15
23	15.8	79.0	5000	4	US-09-791-105B-1
24	15.8	79.0	37792	4	US-09-949-016-12503
25	15.8	79.0	37795	4	US-09-949-016-14263
26	15.8	79.0	46085	4	US-09-949-016-13547
27	15.8	79.0	46085	4	US-09-949-016-13548

28	15.8	79.0	114793	4	US-10-148-806-3	Sequence 3, Appli
29	15.8	79.0	118999	3	US-09-791-105B-32	Sequence 32, Appli
30	15.8	79.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
31	15.8	79.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
32	15.4	77.0	40546	4	US-09-949-016-12847	Sequence 12847, A
33	15.4	77.0	40546	4	US-09-949-016-12915	Sequence 12915, A
34	15.4	77.0	70014	4	US-09-949-016-17110	Sequence 17110, A
35	15.4	77.0	155019	4	US-09-949-016-16029	Sequence 16029, A
36	15.2	76.0	25	4	US-09-396-196G-69161	Sequence 69161, A
37	15.2	76.0	499	4	US-09-513-999C-809	Sequence 809, App
38	15.2	76.0	499	4	US-09-270-767-1373	Sequence 1973, App
39	15.2	76.0	526	4	US-09-270-767-17255	Sequence 17255, A
40	15.2	76.0	597	4	US-09-270-767-5447	Sequence 5447, Ap
41	15.2	76.0	597	4	US-09-270-767-20729	Sequence 20729, A
42	15.2	76.0	601	4	US-09-949-016-122359	Sequence 122359, A
43	15.2	76.0	601	4	US-09-949-016-122360	Sequence 122360, A
44	15.2	76.0	601	4	US-09-949-016-167564	Sequence 167564, A
45	15.2	76.0	601	4	US-09-949-016-167671	Sequence 167671, A
46	15.2	76.0	606	4	US-09-252-991A-607	Sequence 607, App
47	15.2	76.0	834	4	US-09-879-792-28	Sequence 28, Appl
48	15.2	76.0	1164	4	US-09-902-540-8443	Sequence 8443, Ap
49	15.2	76.0	1182	4	US-09-252-991A-520	Sequence 520, App
50	15.2	76.0	1219	4	US-09-270-767-23815	Sequence 23815, A
51	15.2	76.0	1230	4	US-09-879-792-35	Sequence 35, Appl
52	15.2	76.0	1299	4	US-09-252-991A-496	Sequence 496, App
53	15.2	76.0	1341	4	US-10-177-661-3	Sequence 1, Appli
54	15.2	76.0	1434	4	US-10-177-661-1	Sequence 1, Appli
55	15.2	76.0	1551	4	US-09-902-540-6712	Sequence 6712, Ap
56	15.2	76.0	1731	4	US-09-270-767-13778	Sequence 13778, A
57	15.2	76.0	1748	4	US-09-879-792-11	Sequence 11, Appl
58	15.2	76.0	2640	4	US-09-270-767-504	Sequence 504, App
59	15.2	76.0	2640	4	US-09-270-767-15786	Sequence 15786, A
60	15.2	76.0	2643	3	US-09-399-913-56	Sequence 56, Appl
61	15.2	76.0	2643	4	US-09-350-614-55	Sequence 55, Appl
62	15.2	76.0	4019	4	US-09-902-540-535	Sequence 535, App
63	15.2	76.0	7515	4	US-09-902-540-888	Sequence 888, App
64	15.2	76.0	16365	4	US-09-949-016-16970	Sequence 16970, A
65	15.2	76.0	20378	4	US-09-949-016-17281	Sequence 17281, A
66	15.2	76.0	36171	4	US-09-949-016-13876	Sequence 13876, A
67	15.2	76.0	36274	4	US-09-949-016-12389	Sequence 12389, A
68	15.2	76.0	74730	4	US-09-949-016-15189	Sequence 15189, A
69	15.2	76.0	78157	4	US-09-949-016-16466	Sequence 16466, A
70	15.2	76.0	78157	4	US-09-949-016-16467	Sequence 16467, A
71	15.2	76.0	149543	4	US-09-949-016-15947	Sequence 15947, A
72	15.2	76.0	177669	4	US-09-949-016-13713	Sequence 13713, A
73	14.8	74.0	295	3	US-09-018-584A-8	Sequence 8, Appli
74	14.8	74.0	295	4	US-09-784-423-8	Sequence 8, Appli
75	14.8	74.0	491	4	US-09-513-999C-11544	Sequence 11544, A
76	14.8	74.0	601	4	US-09-949-016-23763	Sequence 23763, A
77	14.8	74.0	601	4	US-09-949-016-182417	Sequence 182417, A
78	14.8	74.0	601	4	US-09-949-016-1861134	Sequence 1861134, A
79	14.8	74.0	774	4	US-09-252-991A-9809	Sequence 9809, Ap
80	14.8	74.0	1089	4	US-09-252-991A-9990	Sequence 9990, Ap
81	14.8	74.0	1254	3	US-08-642-274D-49	Sequence 49, Appl
82	14.8	74.0	1254	3	US-08-952-014C-49	Sequence 49, Appl
83	14.8	74.0	3396	4	US-09-614-221A-530	Sequence 530, App
84	14.8	74.0	5912	2	US-08-629-001A-1	Sequence 1, Appli
85	14.8	74.0	5912	3	US-08-642-274D-1	Sequence 1, Appli
86	14.8	74.0	5912	3	US-08-952-127-1	Sequence 1, Appli
87	14.8	74.0	5912	3	US-08-952-014C-1	Sequence 1, Appli
88	14.8	74.0	6525	1	US-08-493-092-3	Sequence 3, Appli
89	14.8	74.0	6525	1	US-08-508-836A-3	Sequence 3, Appli
90	14.8	74.0	7408	1	US-08-441-832-1	Sequence 1, Appli
91	14.8	74.0	7410	1	US-08-493-092-1	Sequence 1, Appli
92	14.8	74.0	7410	1	US-08-508-836A-1	Sequence 1, Appli
93	14.8	74.0	9171	2	US-08-629-001A-2	Sequence 2, Appli
94	14.8	74.0	9171	3	US-08-642-274D-2	Sequence 2, Appli
95	14.8	74.0	9171	3	US-08-952-127-2	Sequence 2, Appli
96	14.8	74.0	9171	3	US-08-952-014C-2	Sequence 2, Appli
97	14.8	74.0	9196	4	US-08-984-090-1	Sequence 1, Appli
98	14.8	74.0	9385	2	US-08-874-266-1	Sequence 1, Appli
99	14.8	74.0	9385	3	US-09-360-416-1	Sequence 1, Appli
100	14.8	74.0	9389	4	US-09-949-016-4409	Sequence 4409, Ap

C 101	14.8	74.0	9870	1	US-08-508-836A-9	Sequence 9, Appl1	C 174	14.2	71.0	601	4	US-09-949-016-99880	Sequence 99880, A
C 102	14.8	74.0	16198	4	US-09-949-016-11854	Sequence 11854, A	C 175	14.2	71.0	601	4	US-09-949-016-100145	Sequence 100145, A
C 103	14.8	74.0	16198	4	US-09-949-016-11854	Sequence 11854, A	C 176	14.2	71.0	601	4	US-09-949-016-100146	Sequence 100146, A
C 104	14.8	74.0	16198	4	US-09-949-016-15973	Sequence 15973, A	C 177	14.2	71.0	601	4	US-09-949-016-100411	Sequence 100411, A
C 105	14.8	74.0	28198	4	US-09-949-016-15973	Sequence 15973, A	C 178	14.2	71.0	601	4	US-09-949-016-100412	Sequence 100412, A
C 106	14.8	74.0	30310	4	US-09-657-346A-96	Sequence 96, Appl1	C 179	14.2	71.0	601	4	US-09-949-016-100677	Sequence 100677, A
C 107	14.8	74.0	56131	4	US-09-949-016-12944	Sequence 12944, A	C 180	14.2	71.0	601	4	US-09-949-016-100678	Sequence 100678, A
C 108	14.8	74.0	136058	4	US-09-949-016-12565	Sequence 12565, A	C 181	14.2	71.0	601	4	US-09-949-016-100981	Sequence 100981, A
C 109	14.8	74.0	136480	4	US-09-949-016-17064	Sequence 17064, A	C 182	14.2	71.0	601	4	US-09-949-016-100982	Sequence 100982, A
C 110	14.8	74.0	146401	4	US-09-949-016-16151	Sequence 16151, A	C 183	14.2	71.0	601	4	US-09-949-016-101247	Sequence 101247, A
C 111	14.8	74.0	153642	4	US-09-949-016-12174	Sequence 12174, A	C 184	14.2	71.0	601	4	US-09-949-016-101248	Sequence 101248, A
C 112	14.8	74.0	153643	4	US-09-949-016-15635	Sequence 15635, A	C 185	14.2	71.0	601	4	US-09-949-016-101513	Sequence 101513, A
C 113	14.8	74.0	202111	4	US-09-949-016-13877	Sequence 13877, A	C 186	14.2	71.0	601	4	US-09-949-016-101514	Sequence 101514, A
C 114	14.8	74.0	254366	4	US-09-822-871-3	Sequence 3, Appl1	C 187	14.2	71.0	601	4	US-09-949-016-101779	Sequence 101779, A
C 115	14.4	72.0	365	4	US-09-621-976-12759	Sequence 12759, A	C 188	14.2	71.0	601	4	US-09-949-016-101780	Sequence 101780, A
C 116	14.4	72.0	601	4	US-09-949-016-47639	Sequence 47639, A	C 189	14.2	71.0	601	4	US-09-949-016-109883	Sequence 109883, A
C 117	14.4	72.0	601	4	US-09-949-016-178213	Sequence 178213, A	C 190	14.2	71.0	601	4	US-09-949-016-115255	Sequence 115255, A
C 118	14.4	72.0	987	4	US-10-061-943A-4	Sequence 4, Appl1	C 191	14.2	71.0	601	4	US-09-949-016-115256	Sequence 115256, A
C 119	14.4	72.0	1441	4	US-09-220-132-39	Sequence 39, Appl1	C 192	14.2	71.0	601	4	US-09-949-016-117714	Sequence 117714, A
C 120	14.4	72.0	4441	4	US-09-949-016-15325	Sequence 15325, A	C 193	14.2	71.0	601	4	US-09-949-016-117731	Sequence 117731, A
C 121	14.4	72.0	5654	4	US-09-573-080A-41	Sequence 41, Appl1	C 194	14.2	71.0	601	4	US-09-949-016-117732	Sequence 117732, A
C 122	14.4	72.0	58782	4	US-09-949-016-16851	Sequence 16851, A	C 195	14.2	71.0	601	4	US-09-949-016-117733	Sequence 117733, A
C 123	14.4	72.0	76472	4	US-09-949-016-15896	Sequence 15896, A	C 196	14.2	71.0	601	4	US-09-949-016-139851	Sequence 139851, A
C 124	14.4	72.0	92276	4	US-09-949-016-12166	Sequence 12166, A	C 197	14.2	71.0	601	4	US-09-949-016-139868	Sequence 139868, A
C 125	14.4	72.0	112239	4	US-09-949-016-13144	Sequence 13144, A	C 198	14.2	71.0	601	4	US-09-949-016-139869	Sequence 139869, A
C 126	14.4	72.0	1830121	4	US-09-557-884-1	Sequence 1, Appl1	C 199	14.2	71.0	601	4	US-09-949-016-139870	Sequence 139870, A
C 127	14.4	72.0	1830121	4	US-09-643-990A-1	Sequence 1, Appl1	C 200	14.2	71.0	601	4	US-09-949-016-168930	Sequence 168930, A
C 128	14.2	71.0	30	3	US-09-052-469-17	Sequence 17, Appl1	C 201	14.2	71.0	601	4	US-09-949-016-168932	Sequence 168932, A
C 129	14.2	71.0	30	4	US-08-422-583-17	Sequence 17, Appl1	C 202	14.2	71.0	711	4	US-09-252-991A-12195	Sequence 12195, A
C 130	14.2	71.0	30	4	US-09-052-262-17	Sequence 17, Appl1	C 203	14.2	71.0	714	4	US-09-252-991A-10604	Sequence 10604, A
C 131	14.2	71.0	315	3	US-08-945-424-1	Sequence 1, Appl1	C 204	14.2	71.0	816	4	US-09-252-991A-10249	Sequence 10249, A
C 132	14.2	71.0	330	1	US-08-718-863-1	Sequence 1, Appl1	C 205	14.2	71.0	870	4	US-09-902-540-6646	Sequence 6646, Ap
C 133	14.2	71.0	330	1	US-08-620-467A-7	Sequence 7, Appl1	C 206	14.2	71.0	991	4	US-09-433-554-27	Sequence 27, Appl1
C 134	14.2	71.0	330	1	US-08-348-572-7	Sequence 7, Appl1	C 207	14.2	71.0	1000	1	US-07-960-112B-1	Sequence 1, Appl1
C 135	14.2	71.0	330	3	US-09-041-090B-7	Sequence 7, Appl1	C 208	14.2	71.0	1000	1	US-07-960-112B-3	Sequence 3, Appl1
C 136	14.2	71.0	330	4	US-08-442-001C-7	Sequence 7, Appl1	C 209	14.2	71.0	1000	1	US-08-301-316B-1	Sequence 1, Appl1
C 137	14.2	71.0	356	4	US-09-513-999C-35305	Sequence 35305, A	C 210	14.2	71.0	1000	1	US-08-301-316B-3	Sequence 3, Appl1
C 138	14.2	71.0	420	3	US-08-943-731-198	Sequence 198, App	C 211	14.2	71.0	1000	2	US-08-473-399B-1	Sequence 1, Appl1
C 139	14.2	71.0	601	4	US-09-949-016-21345	Sequence 21345, A	C 212	14.2	71.0	1000	2	US-08-473-399B-3	Sequence 3, Appl1
C 140	14.2	71.0	601	4	US-09-949-016-57870	Sequence 57870, A	C 213	14.2	71.0	1000	3	US-08-853-831-1	Sequence 1, Appl1
C 141	14.2	71.0	601	4	US-09-949-016-57871	Sequence 57871, A	C 214	14.2	71.0	1000	3	US-08-853-831-3	Sequence 3, Appl1
C 142	14.2	71.0	601	4	US-09-949-016-69385	Sequence 69385, A	C 215	14.2	71.0	1000	3	US-08-853-831-3	Sequence 1, Appl1
C 143	14.2	71.0	601	4	US-09-949-016-69386	Sequence 69386, A	C 216	14.2	71.0	1000	4	US-09-510-885-1	Sequence 1, Appl1
C 144	14.2	71.0	601	4	US-09-949-016-74337	Sequence 74337, A	C 217	14.2	71.0	1000	4	US-09-510-885-3	Sequence 3, Appl1
C 145	14.2	71.0	601	4	US-09-949-016-80719	Sequence 80719, A	C 218	14.2	71.0	1000	5	PCT-US93-09774-1	Sequence 3, Appl1
C 146	14.2	71.0	601	4	US-09-949-016-80720	Sequence 80720, A	C 219	14.2	71.0	1000	5	PCT-US93-09774-3	Sequence 3, Appl1
C 147	14.2	71.0	601	4	US-09-949-016-86421	Sequence 86421, A	C 220	14.2	71.0	1149	4	US-09-252-991A-12262	Sequence 1, Appl1
C 148	14.2	71.0	601	4	US-09-949-016-96421	Sequence 96421, A	C 221	14.2	71.0	1251	4	US-09-252-991A-12303	Sequence 12303, A
C 149	14.2	71.0	601	4	US-09-949-016-96422	Sequence 96422, A	C 222	14.2	71.0	1278	4	US-09-252-991A-12121	Sequence 12121, A
C 150	14.2	71.0	601	4	US-09-949-016-96687	Sequence 96687, A	C 223	14.2	71.0	1356	1	US-08-143-497-1	Sequence 1, Appl1
C 151	14.2	71.0	601	4	US-09-949-016-96688	Sequence 96688, A	C 224	14.2	71.0	1356	1	US-08-461-666-1	Sequence 1, Appl1
C 152	14.2	71.0	601	4	US-09-949-016-96953	Sequence 96953, A	C 225	14.2	71.0	1356	1	US-08-461-666-1	Sequence 1, Appl1
C 153	14.2	71.0	601	4	US-09-949-016-96954	Sequence 96954, A	C 226	14.2	71.0	1356	1	US-08-463-675-1	Sequence 1, Appl1
C 154	14.2	71.0	601	4	US-09-949-016-97219	Sequence 97219, A	C 227	14.2	71.0	1356	1	US-08-757-300-1	Sequence 1, Appl1
C 155	14.2	71.0	601	4	US-09-949-016-97220	Sequence 97220, A	C 228	14.2	71.0	1356	1	US-08-464-589-1	Sequence 1, Appl1
C 156	14.2	71.0	601	4	US-09-949-016-97485	Sequence 97485, A	C 229	14.2	71.0	1386	1	US-09-252-991A-10462	Sequence 10462, A
C 157	14.2	71.0	601	4	US-09-949-016-97486	Sequence 97486, A	C 230	14.2	71.0	1497	4	US-09-799-451-127	Sequence 127, App
C 158	14.2	71.0	601	4	US-09-949-016-97751	Sequence 97751, A	C 231	14.2	71.0	1668	4	US-09-976-594-336	Sequence 336, App
C 159	14.2	71.0	601	4	US-09-949-016-97752	Sequence 97752, A	C 232	14.2	71.0	1756	3	US-09-166-350-30	Sequence 30, Appl
C 160	14.2	71.0	601	4	US-09-949-016-98017	Sequence 98017, A	C 233	14.2	71.0	1756	3	US-09-166-350-30	Sequence 954, App
C 161	14.2	71.0	601	4	US-09-949-016-98018	Sequence 98018, A	C 234	14.2	71.0	1756	4	US-09-023-655-954	Sequence 5563, Ap
C 162	14.2	71.0	601	4	US-09-949-016-98283	Sequence 98283, A	C 235	14.2	71.0	2613	4	US-09-949-016-5563	Sequence 1154, Ap
C 163	14.2	71.0	601	4	US-09-949-016-98284	Sequence 98284, A	C 236	14.2	71.0	2727	4	US-09-949-016-1154	Sequence 1, Appl1
C 164	14.2	71.0	601	4	US-09-949-016-98549	Sequence 98549, A	C 237	14.2	71.0	2878	1	US-07-903-456-1	Sequence 5, Appl1
C 165	14.2	71.0	601	4	US-09-949-016-98550	Sequence 98550, A	C 238	14.2	71.0	2878	3	US-08-666-221B-5	Sequence 11, Appl
C 166	14.2	71.0	601	4	US-09-949-016-98815	Sequence 98815, A	C 239	14.2	71.0	2878	3	US-08-666-221B-13	Sequence 13, Appl
C 167	14.2	71.0	601	4	US-09-949-016-98816	Sequence 98816, A	C 240	14.2	71.0	2878	3	US-08-249-241-1	Sequence 1, Appl1
C 168	14.2	71.0	601	4	US-09-949-016-99081	Sequence 99081, A	C 241	14.2	71.0	3018	4	US-09-051-961-1	Sequence 1, Appl1
C 169	14.2	71.0	601	4	US-09-949-016-99082	Sequence 99082, A	C 242	14.2	71.0	3131	4	US-09-902-540-513	Sequence 513, App
C 170	14.2	71.0	601	4	US-09-949-016-99347	Sequence 99347, A	C 243	14.2	71.0	3540	3	US-08-842-0879-16	Sequence 16, Appl
C 171	14.2	71.0	601	4	US-09-949-016-99348	Sequence 99348, A	C 244	14.2	71.0	3540	4	US-09-638-857-16	Sequence 16, Appl
C 172	14.2	71.0	601	4	US-09-949-016-99613	Sequence 99613, A	C 245	14.2	71.0	4220	3	US-09-183-846A-11	Sequence 11, Appl
C 173	14.2	71.0	601	4	US-09-949-016-99614	Sequence 99614, A	C 246	14.2	71.0	4220	3	US-08-961-578C-11	Sequence 11, Appl
C 174	14.2	71.0	601	4	US-09-949-016-99880	Sequence 99880, A							



247	14.2	71.0	4220	4	US-09-670-216-11	Sequence 11, Appl	320	14.2	71.0	818128	4	US-09-949-016-14560	Sequence 14560, A
248	14.2	71.0	4262	4	US-09-949-016-2134	Sequence 2134, Ap	321	14.2	71.0	818128	4	US-09-949-016-14561	Sequence 14561, A
249	14.2	71.0	4322	4	US-09-949-016-647	Sequence 647, App	322	14.2	71.0	818128	4	US-09-949-016-14562	Sequence 14562, A
250	14.2	71.0	4608	1	US-07-718-575-11	Sequence 11, Appl	323	14.2	71.0	818128	4	US-09-949-016-14564	Sequence 14564, A
251	14.2	71.0	4608	1	US-08-481-206-11	Sequence 11, Appl	324	14.2	71.0	818128	4	US-09-949-016-14565	Sequence 14565, A
252	14.2	71.0	4608	2	US-08-486-289A-11	Sequence 11, Appl	325	14.2	71.0	818128	4	US-09-949-016-14566	Sequence 14566, A
253	14.2	71.0	7090	4	US-09-714-550-18	Sequence 18, Appl	326	14.2	71.0	818128	4	US-09-949-016-14567	Sequence 14567, A
254	14.2	71.0	7881	2	US-08-751-189-1	Sequence 1, Appl	327	14	70.0	601	4	US-09-949-016-78658	Sequence 78658, A
255	14.2	71.0	7881	3	US-09-060-836-1	Sequence 1, Appl	328	14	70.0	601	4	US-09-949-016-84819	Sequence 84819, A
256	14.2	71.0	7881	3	US-09-184-445-1	Sequence 1, Appl	329	14	70.0	78720	4	US-09-949-016-12710	Sequence 12710, A
257	14.2	71.0	9442	4	US-09-949-016-12571	Sequence 12571, A	330	14	70.0	78720	4	US-09-949-016-12710	Sequence 12710, A
258	14.2	71.0	9443	4	US-09-949-016-15877	Sequence 15877, A	331	14	70.0	91062	4	US-09-949-016-17283	Sequence 17283, A
259	14.2	71.0	13941	4	US-09-799-451-341	Sequence 341, App	332	14	70.0	264665	4	US-09-949-016-13019	Sequence 13019, A
260	14.2	71.0	14148	3	US-09-052-469-7	Sequence 7, Appl	333	14	70.0	321022	4	US-09-949-016-13747	Sequence 13747, A
261	14.2	71.0	14148	3	US-08-422-582-7	Sequence 7, Appl	334	14	70.0	321022	4	US-09-949-016-11852	Sequence 11852, A
262	14.2	71.0	14148	4	US-09-052-262-7	Sequence 7, Appl	335	14	70.0	321022	4	US-09-949-016-14166	Sequence 14166, A
263	14.2	71.0	15449	4	US-09-949-016-13600	Sequence 13600, A	336	14	70.0	784019	4	US-09-949-016-14033	Sequence 14033, A
264	14.2	71.0	15356	4	US-09-949-016-13505	Sequence 13505, A	337	14	70.0	828152	4	US-09-949-016-12777	Sequence 12777, A
265	14.2	71.0	17082	4	US-09-949-016-14893	Sequence 14893, A	338	13.8	69.0	25	4	US-09-396-196G-73500	Sequence 73500, A
266	14.2	71.0	20084	3	US-08-943-731-5	Sequence 5, Appl	339	13.8	69.0	119	4	US-09-513-999C-19383	Sequence 19383, A
267	14.2	71.0	21526	4	US-09-949-016-14685	Sequence 14685, A	340	13.8	69.0	312	4	US-09-270-767-26747	Sequence 26747, A
268	14.2	71.0	31571	1	US-08-323-443B-1	Sequence 1, Appl	341	13.8	69.0	485	5	PCT-US95-05741-10	Sequence 10, Appl
269	14.2	71.0	35871	4	US-09-958-335-2	Sequence 2, Appl	342	13.8	69.0	551	4	US-09-949-016-84434	Sequence 84434, A
270	14.2	71.0	35935	2	US-08-735-609-1	Sequence 1, Appl	343	13.8	69.0	593	4	US-09-270-767-10755	Sequence 10755, A
271	14.2	71.0	35935	2	US-08-735-609-1	Sequence 1, Appl	344	13.8	69.0	601	4	US-09-949-016-48801	Sequence 48801, A
272	14.2	71.0	35935	3	US-08-378-452-43	Sequence 43, Appl	345	13.8	69.0	601	4	US-09-949-016-48802	Sequence 48802, A
273	14.2	71.0	35935	3	US-09-315-372-1	Sequence 1, Appl	346	13.8	69.0	601	4	US-09-949-016-48803	Sequence 48803, A
274	14.2	71.0	35935	3	US-09-244-752-1	Sequence 1, Appl	347	13.8	69.0	601	4	US-09-949-016-48973	Sequence 48973, A
275	14.2	71.0	35935	3	US-09-245-437-1	Sequence 1, Appl	348	13.8	69.0	601	4	US-09-949-016-48974	Sequence 48974, A
276	14.2	71.0	35935	3	US-09-409-670-43	Sequence 43, Appl	349	13.8	69.0	601	4	US-09-949-016-48975	Sequence 48975, A
277	14.2	71.0	35935	3	US-09-562-919-1	Sequence 1, Appl	350	13.8	69.0	601	4	US-09-949-016-49145	Sequence 49145, A
278	14.2	71.0	35978	4	US-09-956-335-1	Sequence 1, Appl	351	13.8	69.0	601	4	US-09-949-016-49146	Sequence 49146, A
279	14.2	71.0	39498	4	US-09-949-016-12410	Sequence 12410, A	352	13.8	69.0	601	4	US-09-949-016-49147	Sequence 49147, A
280	14.2	71.0	39498	4	US-09-949-016-16505	Sequence 16505, A	353	13.8	69.0	601	4	US-09-949-016-49321	Sequence 49321, A
281	14.2	71.0	46253	4	US-09-949-016-11890	Sequence 11890, A	354	13.8	69.0	601	4	US-09-949-016-49322	Sequence 49322, A
282	14.2	71.0	46257	4	US-09-949-016-13711	Sequence 13711, A	355	13.8	69.0	601	4	US-09-949-016-49323	Sequence 49323, A
283	14.2	71.0	49399	4	US-09-949-016-13780	Sequence 13780, A	356	13.8	69.0	601	4	US-09-949-016-61189	Sequence 61189, A
284	14.2	71.0	50186	4	US-09-949-016-14066	Sequence 14066, A	357	13.8	69.0	601	4	US-09-949-016-84433	Sequence 84433, A
285	14.2	71.0	53266	3	US-08-658-136-2	Sequence 2, Appl	358	13.8	69.0	601	4	US-09-949-016-85018	Sequence 85018, A
286	14.2	71.0	53577	3	US-08-658-136-1	Sequence 1, Appl	359	13.8	69.0	601	4	US-09-949-016-85169	Sequence 85169, A
287	14.2	71.0	60276	4	US-09-949-016-15004	Sequence 15004, A	360	13.8	69.0	601	4	US-09-949-016-108602	Sequence 108602, A
288	14.2	71.0	60338	4	US-09-949-016-15694	Sequence 15694, A	361	13.8	69.0	601	4	US-09-949-016-120797	Sequence 120797, A
289	14.2	71.0	84761	4	US-09-949-016-11919	Sequence 11919, A	362	13.8	69.0	601	4	US-09-949-016-163641	Sequence 163641, A
290	14.2	71.0	84763	4	US-09-949-016-13914	Sequence 13914, A	363	13.8	69.0	601	4	US-09-949-016-163642	Sequence 163642, A
291	14.2	71.0	92681	4	US-09-949-016-14772	Sequence 14772, A	364	13.8	69.0	601	4	US-09-949-016-168931	Sequence 168931, A
292	14.2	71.0	109038	4	US-09-949-016-12199	Sequence 12199, A	365	13.8	69.0	601	4	US-09-949-016-182411	Sequence 182411, A
293	14.2	71.0	118955	4	US-09-949-016-17565	Sequence 17565, A	366	13.8	69.0	601	4	US-09-949-016-199885	Sequence 199885, A
294	14.2	71.0	117001	4	US-09-949-016-15684	Sequence 15684, A	367	13.8	69.0	601	4	US-09-949-016-199886	Sequence 199886, A
295	14.2	71.0	126176	4	US-09-949-016-16137	Sequence 16137, A	368	13.8	69.0	613	3	US-09-149-476-15	Sequence 15, Appl
296	14.2	71.0	126176	4	US-09-949-016-16138	Sequence 16138, A	369	13.8	69.0	992	4	US-09-270-767-11212	Sequence 11212, A
297	14.2	71.0	137226	4	US-09-949-016-13763	Sequence 13763, A	370	13.8	69.0	1352	1	US-08-552-142A-10	Sequence 10, Appl
298	14.2	71.0	140925	4	US-09-949-016-11777	Sequence 11777, A	371	13.8	69.0	1408	2	US-08-842-657A-10	Sequence 10, Appl
299	14.2	71.0	140982	4	US-09-949-016-16295	Sequence 16295, A	372	13.8	69.0	1414	1	US-08-222-124-9	Sequence 9, Appl
300	14.2	71.0	162914	4	US-09-949-016-15578	Sequence 15578, A	373	13.8	69.0	1414	2	US-08-842-657A-9	Sequence 9, Appl
301	14.2	71.0	236474	4	US-09-949-016-13418	Sequence 13418, A	374	13.8	69.0	1535	1	US-08-910-973-10	Sequence 10, Appl
302	14.2	71.0	254405	4	US-09-949-016-14381	Sequence 14381, A	375	13.8	69.0	1535	3	US-09-499-227-10	Sequence 10, Appl
303	14.2	71.0	258775	4	US-09-949-016-16435	Sequence 16435, A	376	13.8	69.0	1550	3	US-09-234-332-3	Sequence 3, Appl
304	14.2	71.0	451924	4	US-09-949-016-12896	Sequence 12896, A	377	13.8	69.0	1656	3	US-08-318-039A-2	Sequence 2, Appl
305	14.2	71.0	451925	4	US-09-949-016-17305	Sequence 17305, A	378	13.8	69.0	1656	3	US-08-318-038D-16	Sequence 16, Appl
306	14.2	71.0	818128	4	US-09-949-016-14546	Sequence 14546, A	379	13.8	69.0	1656	3	US-08-227-496C-20	Sequence 20, Appl
307	14.2	71.0	818128	4	US-09-949-016-14547	Sequence 14547, A	380	13.8	69.0	1656	3	US-08-435-568A-2	Sequence 2, Appl
308	14.2	71.0	818128	4	US-09-949-016-14548	Sequence 14548, A	381	13.8	69.0	2400	4	US-09-248-796A-2698	Sequence 2698, Ap
309	14.2	71.0	818128	4	US-09-949-016-14549	Sequence 14549, A	382	13.8	69.0	2518	4	US-09-799-451-827	Sequence 827, App
310	14.2	71.0	818128	4	US-09-949-016-14550	Sequence 14550, A	383	13.8	69.0	2586	4	US-10-101-464A-883	Sequence 883, App
311	14.2	71.0	818128	4	US-09-949-016-14551	Sequence 14551, A	384	13.8	69.0	2986	3	US-09-062-416-1	Sequence 1, Appl
312	14.2	71.0	818128	4	US-09-949-016-14552	Sequence 14552, A	385	13.8	69.0	3003	4	US-09-023-655-1115	Sequence 1115, Ap
313	14.2	71.0	818128	4	US-09-949-016-14553	Sequence 14553, A	386	13.8	69.0	3016	2	US-08-344-155C-97	Sequence 97, Appl
314	14.2	71.0	818128	4	US-09-949-016-14554	Sequence 14554, A	387	13.8	69.0	3017	3	US-09-009-490A-86	Sequence 86, Appl
315	14.2	71.0	818128	4	US-09-949-016-14555	Sequence 14555, A	388	13.8	69.0	3024	6	5284931-1	Patent No. 5284931
316	14.2	71.0	818128	4	US-09-949-016-14556	Sequence 14556, A	389	13.8	69.0	3024	6	5284931-1	Patent No. 5284931
317	14.2	71.0	818128	4	US-09-949-016-14557	Sequence 14557, A	390	13.8	69.0	3100	4	US-09-741-238-15	Sequence 15, Appl
318	14.2	71.0	818128	4	US-09-949-016-14558	Sequence 14558, A	391	13.8	69.0	3226	3	US-08-539-205A-5	Sequence 5, Appl
319	14.2	71.0	818128	4	US-09-949-016-14559	Sequence 14559, A	392	13.8	69.0	3226	4	US-09-392-163A-5	Sequence 5, Appl

C 393	13.8	69.0	4899	4	US-09-799-451-714	Sequence 714, Appl	466	13.8	69.0	524032	4	US-09-949-016-16928	Sequence 16928, A
C 394	13.8	69.0	7266	4	US-09-949-016-16428	Sequence 16428, A	C 467	13.8	69.0	524032	4	US-09-949-016-16928	Sequence 16928, A
C 395	13.8	69.0	7386	4	US-09-949-016-13287	Sequence 13287, A	C 468	13.8	69.0	524032	4	US-09-949-016-16929	Sequence 16929, A
C 396	13.8	69.0	8396	3	US-09-328-174A-1	Sequence 1, Appl1	C 469	13.8	69.0	524032	4	US-09-949-016-16929	Sequence 16929, A
C 397	13.8	69.0	8409	3	US-09-167-681-37	Sequence 37, Appl	C 470	13.8	69.0	524032	4	US-09-949-016-16930	Sequence 16930, A
C 398	13.8	69.0	13634	4	US-09-949-016-16376	Sequence 16376, A	C 471	13.8	69.0	524032	4	US-09-949-016-16930	Sequence 16930, A
C 399	13.8	69.0	13906	4	US-09-949-016-14730	Sequence 14730, A	C 472	13.8	69.0	524032	4	US-09-949-016-16931	Sequence 16931, A
C 400	13.8	69.0	14340	4	US-09-949-016-16972	Sequence 16972, A	C 473	13.8	69.0	524032	4	US-09-949-016-16931	Sequence 16931, A
C 401	13.8	69.0	16883	4	US-09-949-016-14694	Sequence 14694, A	C 474	13.8	69.0	529885	4	US-09-949-016-14340	Sequence 14340, A
C 402	13.8	69.0	18302	4	US-09-949-016-13163	Sequence 13163, A	C 475	13.8	69.0	529885	4	US-09-949-016-14340	Sequence 14340, A
C 403	13.8	69.0	22120	4	US-09-949-016-16637	Sequence 16637, A	C 476	13.8	69.0	529885	4	US-09-949-016-14341	Sequence 14341, A
C 404	13.8	69.0	23578	4	US-09-949-016-14160	Sequence 14160, A	C 477	13.8	69.0	529885	4	US-09-949-016-14341	Sequence 14341, A
C 405	13.8	69.0	28555	4	US-09-949-016-13106	Sequence 13106, A	C 478	13.8	69.0	529885	4	US-09-949-016-14342	Sequence 14342, A
C 406	13.8	69.0	31469	4	US-09-949-016-13722	Sequence 13722, A	C 479	13.8	69.0	529885	4	US-09-949-016-14342	Sequence 14342, A
C 407	13.8	69.0	36307	4	US-09-949-016-17372	Sequence 17372, A	C 480	13.8	69.0	529885	4	US-09-949-016-14343	Sequence 14343, A
C 408	13.8	69.0	40037	4	US-09-949-016-12715	Sequence 12715, A	C 481	13.8	69.0	529885	4	US-09-949-016-14343	Sequence 14343, A
C 409	13.8	69.0	42000	4	US-10-081-563-25	Sequence 25, Appl	C 482	13.8	69.0	529885	4	US-09-949-016-14344	Sequence 14344, A
C 410	13.8	69.0	43435	4	US-09-949-016-13909	Sequence 12909, A	C 483	13.8	69.0	529885	4	US-09-949-016-14344	Sequence 14344, A
C 411	13.8	69.0	44821	4	US-09-949-016-13764	Sequence 13764, A	C 484	13.8	69.0	529885	4	US-09-949-016-14345	Sequence 14345, A
C 412	13.8	69.0	54420	4	US-09-949-016-15422	Sequence 15422, A	C 485	13.8	69.0	529885	4	US-09-949-016-14345	Sequence 14345, A
C 413	13.8	69.0	54779	4	US-09-949-001-27	Sequence 27, Appl	C 486	13.8	69.0	529885	4	US-09-949-016-14346	Sequence 14346, A
C 414	13.8	69.0	54780	4	US-09-949-001-39	Sequence 39, Appl	C 487	13.8	69.0	529885	4	US-09-949-016-14346	Sequence 14346, A
C 415	13.8	69.0	57280	4	US-09-949-016-11796	Sequence 11796, A	C 488	13.8	69.0	529885	4	US-09-949-016-14347	Sequence 14347, A
C 416	13.8	69.0	57280	4	US-09-949-016-12843	Sequence 12843, A	C 489	13.8	69.0	529885	4	US-09-949-016-14347	Sequence 14347, A
C 417	13.8	69.0	57280	4	US-09-949-016-12844	Sequence 12844, A	C 490	13.6	68.0	51	4	US-09-443-199C-287	Sequence 287, App
C 418	13.8	69.0	57280	4	US-09-949-016-12846	Sequence 12846, A	C 491	13.6	68.0	67	4	US-09-474-699-7	Sequence 7, Appl1
C 419	13.8	69.0	57280	4	US-09-949-016-13542	Sequence 13542, A	C 492	13.6	68.0	67	4	US-09-814-351-12	Sequence 12, Appl1
C 420	13.8	69.0	57280	4	US-09-949-016-13543	Sequence 13543, A	C 493	13.6	68.0	135	4	US-09-513-999C-14516	Sequence 14516, A
C 421	13.8	69.0	57280	4	US-09-949-016-13544	Sequence 13544, A	C 494	13.6	68.0	137	4	US-09-513-999C-27272	Sequence 27272, A
C 422	13.8	69.0	57280	4	US-09-949-016-13545	Sequence 13545, A	C 495	13.6	68.0	166	4	US-09-513-999C-32611	Sequence 32611, A
C 423	13.8	69.0	57280	4	US-09-949-016-14633	Sequence 14633, A	C 496	13.6	68.0	226	4	US-09-513-999C-23309	Sequence 23309, A
C 424	13.8	69.0	57280	4	US-09-949-016-14634	Sequence 14634, A	C 497	13.6	68.0	253	4	US-09-513-999C-14344	Sequence 14344, A
C 425	13.8	69.0	57280	4	US-09-949-016-14635	Sequence 14635, A	C 498	13.6	68.0	289	4	US-09-313-294A-1541	Sequence 1541, Ap
C 426	13.8	69.0	57280	4	US-09-949-016-14636	Sequence 14636, A	C 499	13.6	68.0	290	3	US-09-280-116-129	Sequence 129, App
C 427	13.8	69.0	57280	4	US-09-949-016-14637	Sequence 14637, A	C 500	13.6	68.0	290	4	US-09-313-294A-6457	Sequence 6457, Ap
C 428	13.8	69.0	57280	4	US-09-949-016-14638	Sequence 14638, A							
C 429	13.8	69.0	57280	4	US-09-949-016-14639	Sequence 14639, A							
C 430	13.8	69.0	57280	4	US-09-949-016-14640	Sequence 14640, A							
C 431	13.8	69.0	68283	4	US-09-949-016-12261	Sequence 12261, A							
C 432	13.8	69.0	68444	4	US-09-949-016-13968	Sequence 13968, A							
C 433	13.8	69.0	68702	4	US-09-949-016-16328	Sequence 16328, A							
C 434	13.8	69.0	70000	3	US-09-851-896-3	Sequence 3, Appl1							
C 435	13.8	69.0	76221	4	US-09-949-016-12009	Sequence 12009, A							
C 436	13.8	69.0	76221	4	US-09-949-016-17053	Sequence 17053, A							
C 437	13.8	69.0	76399	4	US-09-949-016-16819	Sequence 16819, A							
C 438	13.8	69.0	84495	3	US-09-797-906-3	Sequence 3, Appl1							
C 439	13.8	69.0	85869	4	US-09-949-016-12017	Sequence 12017, A							
C 440	13.8	69.0	85878	4	US-09-949-016-16321	Sequence 16321, A							
C 441	13.8	69.0	87190	4	US-09-949-016-16335	Sequence 16335, A							
C 442	13.8	69.0	88906	4	US-09-949-016-17468	Sequence 17468, A							
C 443	13.8	69.0	93510	4	US-09-949-016-15095	Sequence 15095, A							
C 444	13.8	69.0	100463	4	US-09-949-016-12511	Sequence 12511, A							
C 445	13.8	69.0	100468	4	US-09-949-016-12511	Sequence 12511, A							
C 446	13.8	69.0	109159	4	US-09-949-016-14169	Sequence 14169, A							
C 447	13.8	69.0	109159	4	US-09-949-016-14170	Sequence 14170, A							
C 448	13.8	69.0	117937	4	US-09-949-016-12762	Sequence 12762, A							
C 449	13.8	69.0	117937	4	US-09-949-016-15775	Sequence 15775, A							
C 450	13.8	69.0	118382	4	US-09-949-016-15996	Sequence 15996, A							
C 451	13.8	69.0	118382	4	US-09-949-016-15997	Sequence 15997, A							
C 452	13.8	69.0	121068	4	US-09-949-016-14138	Sequence 14138, A							
C 453	13.8	69.0	126468	4	US-09-949-016-14418	Sequence 14418, A							
C 454	13.8	69.0	131724	4	US-09-949-016-12893	Sequence 12893, A							
C 455	13.8	69.0	175265	4	US-09-949-016-16089	Sequence 16089, A							
C 456	13.8	69.0	197336	4	US-09-949-016-12881	Sequence 12881, A							
C 457	13.8	69.0	197337	4	US-09-949-016-14376	Sequence 14376, A							
C 458	13.8	69.0	234288	4	US-09-949-016-17272	Sequence 17272, A							
C 459	13.8	69.0	251672	4	US-09-949-016-17296	Sequence 17296, A							
C 460	13.8	69.0	251682	4	US-09-949-016-11973	Sequence 11973, A							
C 461	13.8	69.0	251769	4	US-09-949-016-13185	Sequence 13185, A							
C 462	13.8	69.0	251769	4	US-09-949-016-13186	Sequence 13186, A							
C 463	13.8	69.0	266748	4	US-09-949-016-13187	Sequence 13187, A							
C 464	13.8	69.0	266748	4	US-09-949-016-13188	Sequence 13188, A							
C 465	13.8	69.0	392000	4	US-10-027-983-11	Sequence 11, Appl							

## ALIGNMENTS

RESULT 1  
US-09-949-016-5659/c  
; Sequence 5659, Application US/09949016  
; Patent No 6812319  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5659  
; LENGTH: 2087  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5659

Query Match 100.0%; Score 20; DB 4; Length 2087;  
Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 AGGGCGTCTCTAGTAGCAG 20  
|||||  
Db 1134 AGGGCGTCTCTAGTAGCAG 1115

```
RESULT 2
US-08-635-552A-1/c
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635.552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 100.0%; Score 20; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20
Db 1134 AGGGCGTCTCTGAGTAGCAG 1115

RESULT 3
US-09-155-768-1/c
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37 001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117

Query Match 100.0%; Score 20; DB 3; Length 2117;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20
Db 1134 AGGGCGTCTCTGAGTAGCAG 1115

RESULT 4
US-09-949-016-17401/c
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

Query Match 100.0%; Score 20; DB 4; Length 14862;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGCGTCTCTGAGTAGCAG 20
Db 11909 AGGGCGTCTCTGAGTAGCAG 11890

RESULT 5
US-09-513-999C-14085
; Sequence 14085, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14085
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-14085
```

Query Match 84.0%; Score 16.8; DB 4; Length 278;  
Best Local Similarity 90.0%; Pred. No. 48;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
DB 113 AGGGCATCTCTGAGTTGCAG 132

RESULT 6  
US-09-513-999C-2290  
; Sequence 2290, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2290  
; LENGTH: 313  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 103...312  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 146  
; OTHER INFORMATION: y=c or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 169  
; OTHER INFORMATION: w=a or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 284  
; OTHER INFORMATION: w=a or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 267  
; OTHER INFORMATION: k=g or t  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 15  
; OTHER INFORMATION: Xaa=Leu or Pro  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 23  
; OTHER INFORMATION: Xaa=Ser or Thr  
US-09-513-999C-2290

Query Match 84.0%; Score 16.8; DB 4; Length 313;  
Best Local Similarity 90.0%; Pred. No. 48;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
DB 113 AGGGCATCTCTGAGTTGCAG 132

RESULT 7  
US-09-904-196B-2  
; Sequence 2, Application US/09904196B  
; Patent No. 655660  
; GENERAL INFORMATION:

; APPLICANT: NISSEN, TORBEN LAUESGAARD  
; APPLICANT: ANDERSEN, KIM VILBOUR  
; APPLICANT: HANSEN, CHRISTIAN KARSTEN  
; APPLICANT: MIKKELSEN, JAN MOLLER  
; TITLE OF INVENTION: G-CSF CONJUGATES  
; FILE REFERENCE: 31-000700US  
; CURRENT APPLICATION NUMBER: US/09/904,196B  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: US/09/760,008  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 60/176,376  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/189,506  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 60/215,644  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DK PA 2000 00024  
; PRIOR FILING DATE: 2000-01-10  
; PRIOR APPLICATION NUMBER: DK PA 2000 00341  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: DK PA 2000 00943  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
US-09-904-196B-2

Query Match 79.0%; Score 15.8; DB 4; Length 525;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||||| |||||  
DB 173 GGGCGCTCTGAGTAGCTG 191

RESULT 8  
US-09-760-008A-2  
; Sequence 2, Application US/09760008A  
; Patent No. 6646110  
; GENERAL INFORMATION:  
; APPLICANT: NISSEN, TORBEN LAUESGAARD  
; APPLICANT: ANDERSEN, KIM VILBOUR  
; APPLICANT: HANSEN, CHRISTIAN KARSTEN  
; APPLICANT: MIKKELSEN, JAN MOLLER  
; TITLE OF INVENTION: G-CSF CONJUGATES  
; FILE REFERENCE: 31-000700US  
; CURRENT APPLICATION NUMBER: US/09/760,008A  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 60/176,376  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/189,506  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 60/215,644  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DK PA 2000 00024  
; PRIOR FILING DATE: 2000-01-10  
; PRIOR APPLICATION NUMBER: DK PA 2000 00341  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: DK PA 2000 00943  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:

; APPLICANT: NISSEN, TORBEN LAUESGAARD  
; APPLICANT: ANDERSEN, KIM VILBOUR  
; APPLICANT: HANSEN, CHRISTIAN KARSTEN  
; APPLICANT: MIKKELSEN, JAN MOLLER  
; TITLE OF INVENTION: G-CSF CONJUGATES  
; FILE REFERENCE: 31-000700US  
; CURRENT APPLICATION NUMBER: US/09/760,008A  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 60/176,376  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/189,506  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 60/215,644  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DK PA 2000 00024  
; PRIOR FILING DATE: 2000-01-10  
; PRIOR APPLICATION NUMBER: DK PA 2000 00341  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: DK PA 2000 00943  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA  
US-09-760-008A-2

Query Match 79.0%; Score 15.8; DB 4; Length 525;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
DB 173 GGGCGCTCTGAGTAGCTG 191

## RESULT 9

US-10-192-294-2  
; Sequence 2, Application US/10192294  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: NISSEN, Torben Lauesgaard  
; TITLE OF INVENTION: G-CSF Conjugates  
; FILE REFERENCE: 0256us310  
; CURRENT APPLICATION NUMBER: US/10/192,294  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: US 09/904,196  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: DK PA 2002 00447  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: DK PA 2002 00708  
; PRIOR FILING DATE: 2002-05-08  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: DNA sequence encoding hG-CSF, with codon usage for E. coli  
US-10-192-294-2

Query Match 79.0%; Score 15.8; DB 4; Length 525;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
DB 173 GGGCGCTCTGAGTAGCTG 191

## RESULT 10

US-09-949-016-62651  
; Sequence 62651, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62651  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-62651

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
DB 533 GGGCGCTCTCTGAGGAGCAG 551

## RESULT 11

US-09-949-016-62652  
; Sequence 62652, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62652  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-62652

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
DB 394 GGGCGCTCTCTGAGGAGCAG 412

## RESULT 12

US-09-949-016-62687  
; Sequence 62687, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62687  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-62687

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20

Db 533 GGGCGTCTCTGAGGACAG 551  
||||| ||||| ||||| |||||

## RESULT 13

US-09-949-016-62688  
; Sequence 62688, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62688  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-62688

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| |||||

Db 394 GGGCGTCTCTGAGGACAG 412

## RESULT 14

US-09-949-016-88843  
; Sequence 88843, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88843  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-88843

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| |||||

Db 357 GGGCGTCTCTGAGGACAG 375

## RESULT 15

US-09-949-016-88844  
; Sequence 88844, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88844  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-88844

Query Match 79.0%; Score 15.8; DB 4; Length 601;  
Best Local Similarity 89.5%; Pred. No. 1.6e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| |||||

Db 376 GGGCGTCTCTGATGAGCAG 394

## RESULT 16

US-09-252-991A-10864/c  
; Sequence 10864, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 10864  
; LENGTH: 882  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-10864

Query Match 79.0%; Score 15.8; DB 4; Length 882;  
Best Local Similarity 89.5%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| |||||

Db 177 GGGCGTCTCGGCTAGCAG 159

## RESULT 17

US-09-902-540-5986/c  
; Sequence 5986, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.



```
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: SINGLE GENE ENCODING AORTIC-SPECIFIC AND STRIATED-SPECIFIC
; TITLE OF INVENTION: MUSCLE CELL ISOFORMS AND USES THEREOF
; FILE REFERENCE: 05433/039001
; CURRENT APPLICATION NUMBER: US/09/303,069A
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: US 09/134,250
; EARLIER FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1806)
US-09-303-069-15

Query Match          79.0%; Score 15.8; DB 3; Length 2614;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGGCGTCTCTGAGTAGCAG 20
Db      1273 GGGCGTCTCTGAGTAGCAG 1255
|||||
RESULT 22
US-09-134-250-15/c
; Sequence 15, Application US/09134250B
; Patent No. 6399753
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Hsieh, Chung-Ming
; TITLE OF INVENTION: SINGLE GENE ENCODING AORTIC-SPECIFIC AND STRIATED-SPECIFIC
; TITLE OF INVENTION: MUSCLE CELL ISOFORMS AND USES THEREOF
; FILE REFERENCE: 05433/038001
; CURRENT APPLICATION NUMBER: US/09/134,250B
; CURRENT FILING DATE: 1998-08-14
; EARLIER APPLICATION NUMBER: US 08/795,868
; EARLIER FILING DATE: 1997-02-06
; EARLIER APPLICATION NUMBER: US 08/494,577
; EARLIER FILING DATE: 1995-06-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(1806)
US-09-134-250-15
```

```
Query Match          79.0%; Score 15.8; DB 3; Length 2614;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GGGCGTCTCTGAGTAGCAG 20
Db      1273 GGGCGTCTCTGAGTAGCAG 1255
|||||
```

```
RESULT 23
US-09-791-105B-1/c
; Sequence 1, Application US/09791105B
; Patent No. 6723508
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie AG
; TITLE OF INVENTION: Method for detecting the presence of at least one single allele d
```

```
; TITLE OF INVENTION: deletion mutant
; FILE REFERENCE: US 09,791,105
; CURRENT APPLICATION NUMBER: US/09/791,105B
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: EP20000103844 20000224
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-791-105B-1
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 5000;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GGGCGTCTCTGAGTAGCAG 20
Db      3623 GGGCGTCTCTGAGTAGCAG 3605
|||||
```

```
RESULT 24
US-09-949-016-12503/c
; Sequence 12503, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12503
; LENGTH: 37792
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12503
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 37792;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 GGGCGTCTCTGAGTAGCAG 20
Db      13159 GGGCGTCTCTGAGTAGCAG 13141
|||||
```

```
RESULT 25
US-09-949-016-14263/c
; Sequence 14263, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```



```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14263
; LENGTH: 37795
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14263

Query Match          79.0%; Score 15.8; DB 4; Length 37795;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
Db 13159 GGGCGTCTCTGAGTAGCAG 13141

RESULT 26
US-09-949-016-13547
; Sequence 13547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13547
; LENGTH: 46085
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(46085)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13547

Query Match          79.0%; Score 15.8; DB 4; Length 46085;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
Db 4900 GGGCGTCTCTGAGTAGCAG 4918

RESULT 27
US-09-949-016-13548
; Sequence 13548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13548
; LENGTH: 46085
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(46085)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13548

Query Match          79.0%; Score 15.8; DB 4; Length 46085;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
Db 4900 GGGCGTCTCTGAGTAGCAG 4918

RESULT 28
US-10-148-806-3
; Sequence 3, Application US/10148806
; Patent No. 6762042
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-148-806-3

Query Match          79.0%; Score 15.8; DB 4; Length 114793;
Best Local Similarity 89.5%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
Db 49994 GGGCGTCTCTGAGTAGCAG 50012

RESULT 29
US-09-791-105B-32/C
; Sequence 32, Application US/09791105B
; Patent No. 6723508
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie AG
; TITLE OF INVENTION: Method for detecting the presence of at least one single allele of
; FILE REFERENCE: US 09/791,105
; CURRENT APPLICATION NUMBER: US/09/791,105B
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: EP20000103844 20000224
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 32
; LENGTH: 118999
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(118999)
; OTHER INFORMATION: Homo sapiens genomic DNA, chromosome 22q11.2, clone KB226F1, Genb
; OTHER INFORMATION: ank accession number AF000351.2 GI:5420485
US-09-791-105B-32

Query Match          79.0%; Score 15.8; DB 4; Length 118999;
Best Local Similarity 89.5%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20
Db 110366 GGGCGTCTCTGAGGAGCAG 110348

RESULT 30
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          79.0%; Score 15.8; DB 3; Length 4403765;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCA 19
Db 495282 AGGCGTCCATGAGTAGCA 495300

RESULT 31
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv

US-09-103-840A-1

Query Match          79.0%; Score 15.8; DB 3; Length 4411529;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCA 19
Db 493840 AGGCGTCCATGAGTAGCA 493858

RESULT 32
US-09-949-016-12847
; Sequence 12847, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12847
; LENGTH: 40546
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-12847

Query Match          77.0%; Score 15.4; DB 4; Length 40546;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GGCCTCTCTGAGTAGCA 19
Db 29363 GGCCTCTCTGAGTAGCA 29379

RESULT 33
US-09-949-016-12915
; Sequence 12915, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12915
; LENGTH: 40546
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-12915

Query Match          77.0%; Score 15.4; DB 4; Length 40546;
Best Local Similarity 94.1%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Query Match	77.0%	Score 15.4	DB 4	Length 155019
Best Local Similarity	94.1%	Pred. No. 4.5e+02		
Matches	16	Conservative	0	Mismatches 1; Indels 0; Gaps 0;
Qy	3	GGCGTCTCTGAGTAGCA	19	
Db	79372	GGCGTCTCTGAGTAGCA	79356	
RESULT 36				
US-09-396-196G-69161/c				
Sequence 69161, Application US/09396196G				
Patent No. 6821724				
GENERAL INFORMATION:				
APPLICANT: Michael Mittmann				
APPLICANT: David Mack				
APPLICANT: David Lockhart				
APPLICANT: Affymetrix, Inc.				
TITLE OF INVENTION: Methods of Genetic Analysis				
FILE REFERENCE: 3101.1				
CURRENT APPLICATION NUMBER: US/09/396,196G				
CURRENT FILING DATE: 1999-09-15				
PRIOR APPLICATION NUMBER: 60/100,678				
PRIOR FILING DATE: 1998-09-17				
NUMBER OF SEQ ID NOS: 127806				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 69161				
LENGTH: 25				
TYPE: DNA				
ORGANISM: mus musculus				
US-09-396-196G-69161				
Query Match	76.0%	Score 15.2	DB 4	Length 25
Best Local Similarity	85.0%	Pred. No. 2.4e+02		
Matches	17	Conservative	0	Mismatches 3; Indels 0; Gaps 0;
Qy	1	AGGCGTCTCTGAGTAGCAG	20	
Db	21	AGGGATTCTCTGTGTAGCAG	2	
RESULT 37				
US-09-513-999C-809/c				
Sequence 809, Application US/09513999C				
Patent No. 6783961				
GENERAL INFORMATION:				
APPLICANT: Dumas Milne Edwards, J.B.				
APPLICANT: Duclert, A.				
APPLICANT: Giordano, J.Y.				
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.				
FILE REFERENCE: 59.US2.REG				
CURRENT APPLICATION NUMBER: US/09/513,999C				
CURRENT FILING DATE: 2000-02-24				
PRIOR APPLICATION NUMBER: US 60/122,487				
PRIOR FILING DATE: 1999-02-26				
NUMBER OF SEQ ID NOS: 36681				
SOFTWARE: Patent.pm				
SEQ ID NO 809				
LENGTH: 499				
TYPE: DNA				
ORGANISM: Homo sapiens				
FEATURE:				
NAME/KEY: CDS				
LOCATION: 170..499				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 165				
OTHER INFORMATION: r=a or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 178				
OTHER INFORMATION: r=a or g				

Query Match	77.0%	Score 15.4	DB 4	Length 155019
Best Local Similarity	94.1%	Pred. No. 4.5e+02		
Matches	16	Conservative	0	Mismatches 1; Indels 0; Gaps 0;
Qy	3	GGCGTCTCTGAGTAGCA	19	
Db	79372	GGCGTCTCTGAGTAGCA	79356	
RESULT 36				
US-09-396-196G-69161/c				
Sequence 69161		Application US/09396196G		
Patent No. 6821724				
GENERAL INFORMATION:				
APPLICANT: Michael Mittmann				
APPLICANT: David Mack				
APPLICANT: David Lockhart				
APPLICANT: Affymetrix, Inc.				
TITLE OF INVENTION: Methods of Genetic Analysis				
FILE REFERENCE: 3101.1				
CURRENT APPLICATION NUMBER: US/09/396,196G				
CURRENT FILING DATE: 1999-09-15				
PRIOR APPLICATION NUMBER: 60/100,678				
PRIOR FILING DATE: 1998-09-17				
NUMBER OF SEQ ID NOS: 127806				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 69161				
LENGTH: 25				
TYPE: DNA				
ORGANISM: mus musculus				
US-09-396-196G-69161				
Query Match	76.0%	Score 15.2	DB 4	Length 25
Best Local Similarity	85.0%	Pred. No. 2.4e+02		
Matches	17	Conservative	0	Mismatches 3; Indels 0; Gaps 0;
Qy	1	AGGCGTCTCTGAGTAGCAG	20	
Db	21	AGGGATTCTCTGTGTAGCAG	2	
RESULT 37				
US-09-513-999C-809/c				
Sequence 809		Application US/09513999C		
Patent No. 6783961				
GENERAL INFORMATION:				
APPLICANT: Dumas Milne Edwards, J.B.				
APPLICANT: Duclert, A.				
APPLICANT: Giordano, J.Y.				
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.				
FILE REFERENCE: 59.US2.REG				
CURRENT APPLICATION NUMBER: US/09/513,999C				
CURRENT FILING DATE: 2000-02-24				
PRIOR APPLICATION NUMBER: US 60/122,487				
PRIOR FILING DATE: 1999-02-26				
NUMBER OF SEQ ID NOS: 36681				
SOFTWARE: Patent.pm				
SEQ ID NO 809				
LENGTH: 499				
TYPE: DNA				
ORGANISM: Homo sapiens				
FEATURE:				
NAME/KEY: CDS				
LOCATION: 170..499				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 165				
OTHER INFORMATION: r=a or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 178				
OTHER INFORMATION: r=a or g				

Query Match	77.0%	Score 15.4	DB 4	Length 155019
Best Local Similarity	94.1%	Pred. No. 4.5e+02		
Matches	16	Conservative	0	Mismatches 1; Indels 0; Gaps 0;
Qy	3	GGCGTCTCTGAGTAGCA	19	
Db	79372	GGCGTCTCTGAGTAGCA	79356	
RESULT 36				
US-09-396-196G-69161/c				
Sequence 69161, Application US/09396196G				
Patent No. 6821724				
GENERAL INFORMATION:				
APPLICANT: Michael Mittmann				
APPLICANT: David Mack				
APPLICANT: David Lockhart				
APPLICANT: Affymetrix, Inc.				
TITLE OF INVENTION: Methods of Genetic Analysis				
FILE REFERENCE: 3101.1				
CURRENT APPLICATION NUMBER: US/09/396,196G				
CURRENT FILING DATE: 1999-09-15				
PRIOR APPLICATION NUMBER: 60/100,678				
PRIOR FILING DATE: 1998-09-17				
NUMBER OF SEQ ID NOS: 127806				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 69161				
LENGTH: 25				
TYPE: DNA				
ORGANISM: mus musculus				
US-09-396-196G-69161				
Query Match	76.0%	Score 15.2	DB 4	Length 25
Best Local Similarity	85.0%	Pred. No. 2.4e+02		
Matches	17	Conservative	0	Mismatches 3; Indels 0; Gaps 0;
Qy	1	AGGCGTCTCTGAGTAGCAG	20	
Db	21	AGGGATTCTCTGTGTAGCAG	2	
RESULT 37				
US-09-513-999C-809/c				
Sequence 809, Application US/09513999C				
Patent No. 6783961				
GENERAL INFORMATION:				
APPLICANT: Dumas Milne Edwards, J.B.				
APPLICANT: Duclert, A.				
APPLICANT: Giordano, J.Y.				
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.				
FILE REFERENCE: 59.US2.REG				
CURRENT APPLICATION NUMBER: US/09/513,999C				
CURRENT FILING DATE: 2000-02-24				
PRIOR APPLICATION NUMBER: US 60/122,487				
PRIOR FILING DATE: 1999-02-26				
NUMBER OF SEQ ID NOS: 36681				
SOFTWARE: Patent.pm				
SEQ ID NO 809				
LENGTH: 499				
TYPE: DNA				
ORGANISM: Homo sapiens				
FEATURE:				
NAME/KEY: CDS				
LOCATION: 170..499				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 165				
OTHER INFORMATION: r=a or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: 178				
OTHER INFORMATION: r=a or g				

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 184
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 196
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 198
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 214
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 232
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 234
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 238
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 241
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 389
; OTHER INFORMATION: r=a or g
; NAME/KEY: UNSURE
; LOCATION: 10
; OTHER INFORMATION: Xaa=Ala or Asp
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 15
; OTHER INFORMATION: Xaa=His or Gln
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 22
; OTHER INFORMATION: Xaa=Asp or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 74
; OTHER INFORMATION: Xaa=Ala or Thr
US-09-513-999C-809

Query Match          76.0%; Score 15.2; DB 4; Length 499;
Best Local Similarity 85.0%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTCAGTAGCAG 20
Db      84 AGGGCGTCTCTCAGTAGCAG 65

RESULT 38
US-09-270-767-1973
; Sequence 1973, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5447
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5447

Query Match          76.0%; Score 15.2; DB 4; Length 597;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTCAGTAGCAG 20
Db      190 AGGGCAGCTCTCAGTAGCAG 171

RESULT 41
US-09-270-767-20729/c

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1973
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1973

Query Match          76.0%; Score 15.2; DB 4; Length 526;
Best Local Similarity 85.0%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTCAGTAGCAG 20
Db      121 AGGGCGTCTCCGAGCAGCG 140

RESULT 39
US-09-270-767-17255
; Sequence 17255, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17255
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17255

Query Match          76.0%; Score 15.2; DB 4; Length 526;
Best Local Similarity 85.0%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTCAGTAGCAG 20
Db      121 AGGGCGTCTCCGAGCAGCG 140

RESULT 40
US-09-270-767-5447/c
; Sequence 5447, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5447
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5447

Query Match          76.0%; Score 15.2; DB 4; Length 597;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTCAGTAGCAG 20
Db      190 AGGGCAGCTCTCAGTAGCAG 171

RESULT 41
US-09-270-767-20729/c
```

```
; Sequence 20729, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20729
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20729

Query Match          76.0%; Score 15.2; DB 4; Length 597;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 190 AGGGCAGCTCTGAGCAGCAG 171

RESULT 42
US-09-949-016-122359/c
; Sequence 122359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122359
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122359

Query Match          76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 377 AGGGCTACCTGAGTAGCAG 358

RESULT 43
US-09-949-016-122360/c
; Sequence 122360, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122360
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-122360

Query Match          76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 532 ATGGAGTCTCAGAGTAGCAG 513

RESULT 44
US-09-949-016-167564/c
; Sequence 167564, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167564
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167564

Query Match          76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
Db 532 ATGGAGTCTCAGAGTAGCAG 513

RESULT 45
US-09-949-016-167671/c
; Sequence 167671, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167671
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167671
```

; SEQ ID NO 167671  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-167671

Query Match 76.0%; Score 15.2; DB 4; Length 601;  
Best Local Similarity 85.0%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
||| ||||| ||||| |||||  
Db 532 ATGGAGTCTCAGAGTAGCAG 513

RESULT 46  
US-09-252-991A-607  
; Sequence 607, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 607  
; LENGTH: 606  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-607

Query Match 76.0%; Score 15.2; DB 4; Length 606;  
Best Local Similarity 85.0%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
||| ||||| ||||| |||||  
Db 503 AGGCGCGCTGAGCAGCAG 522

RESULT 47  
US-09-879-792-28/c  
; Sequence 28, Application US/09879792  
; Patent No. 6734006  
; GENERAL INFORMATION:  
; APPLICANT: Xiao, Yonghong  
; APPLICANT: Gedrich, Richard  
; TITLE OF INVENTION: Regulation of Human Transmembrane Sertine  
; FILE REFERENCE: 02973.00035  
; CURRENT APPLICATION NUMBER: US/09/879,792  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/211,224  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US 60/283,353  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/283,648  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT (Docket No. 6734006 LIO-81-WO)  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 834  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-879-792-28

Query Match 76.0%; Score 15.2; DB 4; Length 834;  
Best Local Similarity 85.0%; Pred. No. 3.4e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
||| ||||| ||||| |||||  
Db 119 AGGTCTTCTCTGAGTAGGAG 100

RESULT 48  
US-09-902-540-8443/c  
; Sequence 8443, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 8443  
; LENGTH: 1164  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-8443

Query Match 76.0%; Score 15.2; DB 4; Length 1164;  
Best Local Similarity 85.0%; Pred. No. 3.5e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
||| ||||| ||||| |||||  
Db 104 AGGCGCTCTGAGCAGCAG 85

RESULT 49  
US-09-252-991A-520/c  
; Sequence 520, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 520  
; LENGTH: 1182  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-520

Query Match 76.0%; Score 15.2; DB 4; Length 1182;  
Best Local Similarity 85.0%; Pred. No. 3.5e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
||| ||||| ||||| |||||  
Db 200 AGGCGCGCTGAGCAGCAG 181

RESULT 50

```

US-09-270-767-29815
; Sequence 29815, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29815
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29815

Query Match      76.0%; Score 15.2; DB 4; Length 1219;
Best Local Similarity 85.0%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTGAGTAGCAG 20
Db      627 AGGGCGTCTCTGAGTAGCAG 646

```

Search completed: March 14, 2005, 20:48:00  
Job time : 149 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 20:46:15 ; Search time 328 Seconds  
(without alignments)  
362.715 Million cell updates/sec

Title: US-10-672-399-10

Perfect score: 20

Sequence: 1 agggcgctctgtagtagcag 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	18	US-10-672-399-10
2	20	100.0	1071	18	US-10-672-399-3
3	20	100.0	1083	18	US-10-672-399-5
4	20	100.0	1737	18	US-10-672-399-1
5	20	100.0	2116	13	US-10-042-523-1
6	17.4	87.0	1752	10	US-09-902-939-1
7	16.8	84.0	188	18	US-10-425-115-88239
8	16.8	84.0	239	16	US-10-023-386-22234
9	16.8	84.0	595	16	US-10-029-386-8534
10	16.8	84.0	80105	18	US-10-322-281-223
11	16.4	82.0	300	14	US-10-076-555-218
					Sequence 10, Appl
					Sequence 3, Appl
					Sequence 5, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 1, Appl
					Sequence 88239, A
					Sequence 22334, A
					Sequence 8534, Ap
					Sequence 223, App
					Sequence 218, App

Sequence 842, App  
Sequence 3940, Ap  
Sequence 1094, Ap  
Sequence 256, App  
Sequence 795, App  
Sequence 6203, Ap  
Sequence 28588, A  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 27903, A  
Sequence 142002, A  
Sequence 169596, A  
Sequence 142002, A  
Sequence 169596, A  
Sequence 116504, A  
Sequence 116504, A  
Sequence 16915, A  
Sequence 34349, A  
Sequence 110517, A  
Sequence 67263, A  
Sequence 111979, A  
Sequence 111979, A  
Sequence 15, Appl  
Sequence 1, Appli  
Sequence 98, Appl  
Sequence 87, Appl  
Sequence 87, Appl  
Sequence 124, App  
Sequence 68, Appl  
Sequence 3, Appli  
Sequence 199, App  
Sequence 17071, A  
Sequence 1507, Ap  
Sequence 242, App  
Sequence 10350, A  
Sequence 290561, A  
Sequence 290561, A  
Sequence 143571, A  
Sequence 143571, A  
Sequence 620, App  
Sequence 148, App  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 1408, Ap  
Sequence 12, Appl  
Sequence 69161, A  
Sequence 102822, A  
Sequence 31331, A  
Sequence 92422, A  
Sequence 2542, Ap  
Sequence 41159, A  
Sequence 41159, A  
Sequence 41159, A  
Sequence 532, App  
Sequence 6722, Ap  
Sequence 86867, A  
Sequence 303658, A  
Sequence 86867, A  
Sequence 303658, A  
Sequence 23748, A  
Sequence 20665, A  
Sequence 380, App  
Sequence 380, App  
Sequence 14834, A  
Sequence 127748, A  
Sequence 127748, A  
Sequence 127750, A  
Sequence 127751, A  
Sequence 127748, A

C 85	15.2	76.0	645	17	US-10-027-632-127749	Sequence 127749,	158	15.2	76.0	1980090	18	US-10-719-993-6815	Sequence 6815, A
C 86	15.2	76.0	645	17	US-10-027-632-127750	Sequence 127750,	159	15.2	76.0	1980090	19	US-10-741-600-17676	Sequence 17676, A
C 87	15.2	76.0	645	17	US-10-027-632-127751	Sequence 127751,	C 160	15	75.0	456	10	US-09-918-995-20598	Sequence 20598, A
C 88	15.2	76.0	684	10	US-09-972-656-83	Sequence 83, Appl	C 161	15	75.0	552	16	US-10-029-386-11879	Sequence 11879, A
C 89	15.2	76.0	741	18	US-10-437-963-822	Sequence 822, App	C 162	15	75.0	1379	17	US-10-291-265-36	Sequence 36, Appl
C 90	15.2	76.0	834	9	US-09-879-792-28	Sequence 28, Appl	C 163	15	75.0	11876	18	US-10-723-860-1800	Sequence 1800, Ap
C 91	15.2	76.0	834	18	US-10-806-370-28	Sequence 28, Appl	C 164	15	75.0	11919	18	US-10-723-860-6158	Sequence 6158, Ap
C 92	15.2	76.0	842	18	US-10-425-115-158723	Sequence 158723,	C 165	15	75.0	65464	9	US-09-859-888-3	Sequence 3, Appli
C 93	15.2	76.0	995	13	US-10-021-509-4	Sequence 4, Appli	C 166	15	75.0	65464	18	US-10-476-543-3	Sequence 3, Appli
C 94	15.2	76.0	995	13	US-10-021-509-12	Sequence 12, Appl	C 167	15	75.0	200418	13	US-10-087-192-568	Sequence 568, App
C 95	15.2	76.0	996	18	US-10-425-115-96589	Sequence 96589, A	C 168	14.8	74.0	25	19	US-10-719-900-462907	Sequence 462907,
C 96	15.2	76.0	1078	17	US-09-898-837A-16	Sequence 16, Appl	C 169	14.8	74.0	261	17	US-10-424-599-67106	Sequence 67106, A
C 97	15.2	76.0	1101	17	US-10-428-275-109	Sequence 109, App	C 170	14.8	74.0	295	9	US-09-784-423-8	Sequence 8, Appli
C 98	15.2	76.0	1110	17	US-10-428-275-113	Sequence 113, App	C 171	14.8	74.0	323	18	US-10-425-115-118330	Sequence 118330,
C 99	15.2	76.0	1146	17	US-10-428-275-133	Sequence 133, App	C 172	14.8	74.0	333	18	US-10-437-963-50208	Sequence 50208, A
C 100	15.2	76.0	1203	17	US-10-428-275-111	Sequence 111, App	C 173	14.8	74.0	376	17	US-10-424-599-26657	Sequence 26657, A
C 101	15.2	76.0	1203	17	US-10-428-275-115	Sequence 115, App	C 174	14.8	74.0	411	10	US-09-918-995-4283	Sequence 4283, Ap
C 102	15.2	76.0	1203	17	US-10-428-275-117	Sequence 117, App	C 175	14.8	74.0	422	18	US-10-914-016-121	Sequence 121, App
C 103	15.2	76.0	1218	17	US-10-428-275-125	Sequence 125, App	C 176	14.8	74.0	455	10	US-09-918-995-27588	Sequence 27588, A
C 104	15.2	76.0	1218	17	US-10-428-275-137	Sequence 137, App	C 177	14.8	74.0	467	9	US-09-864-761-1521	Sequence 1521, Ap
C 105	15.2	76.0	1230	9	US-09-879-792-35	Sequence 35, Appl	C 178	14.8	74.0	526	13	US-10-027-632-274696	Sequence 274696,
C 106	15.2	76.0	1230	18	US-10-806-370-35	Sequence 35, Appl	C 179	14.8	74.0	526	17	US-10-027-632-274696	Sequence 274696,
C 107	15.2	76.0	1314	10	US-09-898-837A-14	Sequence 14, Appl	C 180	14.8	74.0	547	17	US-10-424-599-70440	Sequence 70440, A
C 108	15.2	76.0	1314	17	US-09-898-837A-17	Sequence 17, Appl	C 181	14.8	74.0	552	18	US-10-335-053-245	Sequence 245, App
C 109	15.2	76.0	1314	17	US-10-428-275-131	Sequence 131, App	C 182	14.8	74.0	572	9	US-09-864-761-18280	Sequence 18280, A
C 110	15.2	76.0	1341	14	US-10-177-661-3	Sequence 3, Appli	C 183	14.8	74.0	646	13	US-10-027-632-122441	Sequence 122441,
C 111	15.2	76.0	1419	18	US-10-437-963-61831	Sequence 61831, A	C 184	14.8	74.0	646	17	US-10-027-632-122441	Sequence 122441,
C 112	15.2	76.0	1434	14	US-10-177-661-1	Sequence 1, Appli	C 185	14.8	74.0	707	17	US-10-260-238-3941	Sequence 3941, Ap
C 113	15.2	76.0	1574	17	US-10-274-639-29	Sequence 29, Appl	C 186	14.8	74.0	723	18	US-10-653-047-4558	Sequence 4558, Ap
C 114	15.2	76.0	1574	17	US-10-333-574-29	Sequence 29, Appl	C 187	14.8	74.0	731	18	US-10-437-963-41552	Sequence 41552, A
C 115	15.2	76.0	1589	18	US-10-425-115-101269	Sequence 101269,	C 188	14.8	74.0	785	17	US-10-260-238-4613	Sequence 4613, Ap
C 116	15.2	76.0	1614	9	US-09-888-615-45	Sequence 45, Appl	C 189	14.8	74.0	838	13	US-10-027-632-160660	Sequence 160660,
C 117	15.2	76.0	1626	17	US-10-428-275-121	Sequence 121, App	C 190	14.8	74.0	838	17	US-10-027-632-160660	Sequence 160660,
C 118	15.2	76.0	1626	18	US-10-425-115-86930	Sequence 86930, A	C 191	14.8	74.0	884	14	US-10-076-555-796	Sequence 796, App
C 119	15.2	76.0	1689	17	US-10-156-214A-23	Sequence 23, Appl	C 192	14.8	74.0	1158	10	US-09-372-348-28	Sequence 28, Appl
C 120	15.2	76.0	1707	17	US-10-428-275-129	Sequence 129, App	C 193	14.8	74.0	1168	9	US-09-778-844-43	Sequence 43, Appl
C 121	15.2	76.0	1708	17	US-10-428-275-147	Sequence 147, App	C 194	14.8	74.0	1254	10	US-09-825-476-49	Sequence 49, Appl
C 122	15.2	76.0	1748	9	US-09-879-792-11	Sequence 11, Appl	C 195	14.8	74.0	1414	18	US-10-437-963-825	Sequence 825, App
C 123	15.2	76.0	1748	18	US-10-806-370-11	Sequence 11, Appl	C 196	14.8	74.0	2271	15	US-10-259-165-149	Sequence 149, App
C 124	15.2	76.0	1771	17	US-10-428-275-123	Sequence 123, App	C 197	14.8	74.0	2346	17	US-10-062-674-1878	Sequence 1878, Ap
C 125	15.2	76.0	1779	17	US-10-428-275-119	Sequence 119, App	C 198	14.8	74.0	2455	18	US-10-425-115-148811	Sequence 148811,
C 126	15.2	76.0	2067	17	US-10-156-214A-25	Sequence 25, Appl	C 199	14.8	74.0	2484	17	US-10-108-260A-1543	Sequence 1543, Ap
C 127	15.2	76.0	2310	17	US-10-428-275-127	Sequence 127, App	C 200	14.8	74.0	2674	17	US-10-295-027-655	Sequence 655, App
C 128	15.2	76.0	2393	17	US-10-353-690-99	Sequence 99, Appl	C 201	14.8	74.0	2691	18	US-10-723-860-6540	Sequence 6540, Ap
C 129	15.2	76.0	2410	17	US-10-311-035-35	Sequence 35, Appl	C 202	14.8	74.0	2896	18	US-10-437-963-9614	Sequence 9614, Ap
C 130	15.2	76.0	2432	17	US-10-428-275-103	Sequence 103, App	C 203	14.8	74.0	3012	17	US-10-108-260A-1713	Sequence 1713, Ap
C 131	15.2	76.0	2432	17	US-10-428-275-139	Sequence 139, App	C 204	14.8	74.0	3396	18	US-10-793-639-530	Sequence 530, App
C 132	15.2	76.0	2432	17	US-10-428-275-141	Sequence 141, App	C 205	14.8	74.0	3589	18	US-10-478-914-21	Sequence 21, Appl
C 133	15.2	76.0	2432	17	US-10-428-275-143	Sequence 143, App	C 206	14.8	74.0	5912	10	US-09-825-476-1	Sequence 1, Appli
C 134	15.2	76.0	2643	9	US-09-350-874-56	Sequence 56, Appl	C 207	14.8	74.0	9171	9	US-09-810-993-1	Sequence 1, Appli
C 135	15.2	76.0	2643	13	US-10-106-989-56	Sequence 56, Appl	C 208	14.8	74.0	9171	10	US-09-825-476-2	Sequence 2, Appli
C 136	15.2	76.0	2717	13	US-10-027-632-111854	Sequence 111854,	C 209	14.8	74.0	9171	16	US-10-251-210-1	Sequence 1, Appli
C 137	15.2	76.0	2717	17	US-10-027-632-111854	Sequence 111854,	C 210	14.8	74.0	9196	8	US-08-984-090-1	Sequence 1, Appli
C 138	15.2	76.0	2727	17	US-10-291-172-4	Sequence 4, Appli	C 211	14.8	74.0	9196	15	US-10-361-562-1	Sequence 1, Appli
C 139	15.2	76.0	2727	17	US-10-221-278-4	Sequence 4, Appli	C 212	14.8	74.0	9385	14	US-10-175-225-1	Sequence 1, Appli
C 140	15.2	76.0	2772	16	US-10-119-428-4	Sequence 4, Appli	C 213	14.8	74.0	10948	15	US-10-017-161-2069	Sequence 2069, Ap
C 141	15.2	76.0	3678	17	US-10-425-114-13558	Sequence 13558, A	C 214	14.8	74.0	10948	17	US-10-292-798-1715	Sequence 1715, Ap
C 142	15.2	76.0	4859	8	US-10-425-115-118653	Sequence 118653,	C 215	14.8	74.0	17286	17	US-10-366-345-8	Sequence 8, Appli
C 143	15.2	76.0	5728	9	US-09-917-800A-479	Sequence 479, App	C 216	14.8	74.0	17286	18	US-10-723-860-2345	Sequence 2345, Ap
C 144	15.2	76.0	5728	17	US-10-388-934-73	Sequence 73, Appl	C 217	14.8	74.0	21898	17	US-10-034-650-7	Sequence 7, Appli
C 145	15.2	76.0	7513	15	US-10-017-161-1779	Sequence 161, App	C 218	14.8	74.0	25543	13	US-10-087-192-805	Sequence 805, App
C 146	15.2	76.0	7513	17	US-10-292-798-1435	Sequence 1435, Ap	C 219	14.8	74.0	26345	13	US-10-087-192-1705	Sequence 1705, Ap
C 147	15.2	76.0	15515	9	US-09-822-860-3	Sequence 3, Appli	C 220	14.8	74.0	26345	13	US-10-087-192-1705	Sequence 1705, Ap
C 148	15.2	76.0	25083	11	US-09-997-722-121	Sequence 121, App	C 221	14.8	74.0	28350	17	US-10-085-117-169	Sequence 169, App
C 149	15.2	76.0	25231	10	US-09-764-891-5800	Sequence 5800, Ap	C 222	14.8	74.0	30310	9	US-09-800-631-96	Sequence 96, Appl
C 150	15.2	76.0	44377	17	US-10-085-117-40	Sequence 40, Appl	C 223	14.8	74.0	30310	15	US-10-293-783-96	Sequence 96, Appl
C 151	15.2	76.0	49914	18	US-10-322-281-299	Sequence 299, App	C 224	14.8	74.0	30310	17	US-10-388-263-745	Sequence 745, App
C 152	15.2	76.0	63248	13	US-10-087-192-1879	Sequence 1879, Ap	C						

c 231	14.8	74.0	37135	17	US-10-052-482-151	Sequence 151, App	c 304	14.4	72.0	456	18	US-10-357-930-524	Sequence 524, App
c 232	14.8	74.0	39924	13	US-10-087-192-1993	Sequence 1993, Ap	c 305	14.4	72.0	473	18	US-10-723-860-3375	Sequence 3375, Ap
c 233	14.8	74.0	41522	13	US-10-087-192-1741	Sequence 1741, Ap	c 306	14.4	72.0	505	10	US-09-918-995-20349	Sequence 20349, A
c 234	14.8	74.0	42179	13	US-10-087-192-1969	Sequence 1969, Ap	c 307	14.4	72.0	560	17	US-10-152-319A-370	Sequence 370, App
c 235	14.8	74.0	42339	13	US-10-087-192-991	Sequence 991, App	c 308	14.4	72.0	589	13	US-10-027-632-135590	Sequence 135590,
c 236	14.8	74.0	44990	17	US-10-052-482-217	Sequence 217, App	c 309	14.4	72.0	589	17	US-10-027-632-135590	Sequence 135590,
c 237	14.8	74.0	45315	18	US-10-322-281-785	Sequence 785, App	c 310	14.4	72.0	593	16	US-10-029-386-4162	Sequence 4162, Ap
c 238	14.8	74.0	45606	17	US-10-085-117-253	Sequence 253, App	c 311	14.4	72.0	644	13	US-10-027-632-239018	Sequence 239018,
c 239	14.8	74.0	46137	18	US-10-322-281-353	Sequence 353, App	c 312	14.4	72.0	644	13	US-10-027-632-239018	Sequence 239018,
c 240	14.8	74.0	46137	18	US-10-322-281-353	Sequence 353, App	c 313	14.4	72.0	644	13	US-10-027-632-239018	Sequence 239018,
c 241	14.8	74.0	49502	13	US-10-087-192-1099	Sequence 1099, Ap	c 314	14.4	72.0	644	17	US-10-027-632-239019	Sequence 239019,
c 242	14.8	74.0	52312	18	US-10-322-281-753	Sequence 753, App	c 315	14.4	72.0	665	13	US-10-027-632-198234	Sequence 198234,
c 243	14.8	74.0	52637	18	US-10-367-094-186	Sequence 186, App	c 316	14.4	72.0	665	17	US-10-027-632-198234	Sequence 198234,
c 244	14.8	74.0	54810	18	US-10-417-375-91	Sequence 91, Appl	c 317	14.4	72.0	716	18	US-10-425-115-58096	Sequence 58096, A
c 245	14.8	74.0	55606	17	US-10-085-117-133	Sequence 133, App	c 318	14.4	72.0	756	13	US-10-027-632-147299	Sequence 147299,
c 246	14.8	74.0	62278	18	US-10-322-281-849	Sequence 849, App	c 319	14.4	72.0	756	13	US-10-027-632-147299	Sequence 147299,
c 247	14.8	74.0	65787	17	US-10-052-482-109	Sequence 109, App	c 320	14.4	72.0	756	17	US-10-027-632-147299	Sequence 147299,
c 248	14.8	74.0	66881	18	US-10-322-281-411	Sequence 411, App	c 321	14.4	72.0	756	17	US-10-027-632-147300	Sequence 147300,
c 249	14.8	74.0	67191	11	US-09-997-722-169	Sequence 169, App	c 322	14.4	72.0	759	9	US-09-822-849A-546	Sequence 546, App
c 250	14.8	74.0	67191	16	US-10-105-612-1	Sequence 1, Appli	c 323	14.4	72.0	779	18	US-10-425-115-58094	Sequence 58094, A
c 251	14.8	74.0	68233	17	US-10-034-650-31	Sequence 31, Appl	c 324	14.4	72.0	790	18	US-10-425-115-58094	Sequence 58094, A
c 252	14.8	74.0	73900	17	US-10-085-117-295	Sequence 295, App	c 325	14.4	72.0	872	19	US-10-487-561-19	Sequence 19, Appl
c 253	14.8	74.0	76150	17	US-10-085-117-157	Sequence 157, App	c 326	14.4	72.0	970	18	US-10-723-860-7358	Sequence 7358, Ap
c 254	14.8	74.0	80423	18	US-10-367-094-41	Sequence 41, Appl	c 327	14.4	72.0	987	10	US-09-944-049-11	Sequence 11, Appl
c 255	14.8	74.0	80766	17	US-10-085-117-97	Sequence 97, Appl	c 328	14.4	72.0	987	13	US-10-061-943A-4	Sequence 4, Appli
c 256	14.8	74.0	87415	13	US-10-087-192-1651	Sequence 1651, Ap	c 329	14.4	72.0	987	17	US-10-681-478-4	Sequence 4, Appli
c 257	14.8	74.0	87464	13	US-10-087-192-1411	Sequence 1411, Ap	c 330	14.4	72.0	1058	9	US-09-864-761-16865	Sequence 16865, A
c 258	14.8	74.0	90043	13	US-10-087-192-1141	Sequence 1141, Ap	c 331	14.4	72.0	1419	13	US-10-108-605-140	Sequence 140, App
c 259	14.8	74.0	90351	18	US-10-367-094-166	Sequence 166, App	c 332	14.4	72.0	1601	10	US-09-960-706-465	Sequence 465, App
c 260	14.8	74.0	90351	18	US-10-367-094-166	Sequence 166, App	c 333	14.4	72.0	1601	10	US-09-960-706-465	Sequence 465, App
c 261	14.8	74.0	96389	17	US-10-052-482-181	Sequence 181, App	c 334	14.4	72.0	1601	18	US-10-717-597-116	Sequence 116, App
c 262	14.8	74.0	96597	17	US-10-052-482-226	Sequence 226, App	c 335	14.4	72.0	1672	18	US-10-370-715B-129	Sequence 129, App
c 263	14.8	74.0	97415	18	US-10-322-281-287	Sequence 287, App	c 336	14.4	72.0	2346	17	US-10-205-331-113	Sequence 113, App
c 264	14.8	74.0	100554	13	US-10-087-192-1165	Sequence 1165, Ap	c 337	14.4	72.0	2362	14	US-10-198-846-13946	Sequence 13946, A
c 265	14.8	74.0	102145	18	US-10-331-053-53	Sequence 53, Appl	c 338	14.4	72.0	2505	19	US-10-734-049A-315	Sequence 315, App
c 266	14.8	74.0	114613	13	US-10-087-192-676	Sequence 676, App	c 339	14.4	72.0	2589	18	US-10-425-115-175990	Sequence 175990,
c 267	14.8	74.0	114613	13	US-10-087-192-727	Sequence 727, App	c 340	14.4	72.0	3186	14	US-10-198-846-10881	Sequence 10881, A
c 268	14.8	74.0	115223	13	US-10-322-281-773	Sequence 773, App	c 341	14.4	72.0	3438	17	US-10-369-493A-25413	Sequence 25413, A
c 269	14.8	74.0	123920	13	US-10-087-192-1453	Sequence 1453, Ap	c 342	14.4	72.0	3685	9	US-09-822-830A-331	Sequence 331, App
c 270	14.8	74.0	131217	18	US-10-779-597-1	Sequence 1, Appli	c 343	14.4	72.0	5131	9	US-09-764-877-3800	Sequence 3800, Ap
c 271	14.8	74.0	140152	18	US-10-684-422-66	Sequence 66, Appl	c 344	14.4	72.0	5131	17	US-10-242-515-3800	Sequence 3800, Ap
c 272	14.8	74.0	148497	18	US-10-723-860-1744	Sequence 1744, Ap	c 345	14.4	72.0	5139	9	US-09-764-877-3802	Sequence 3802, Ap
c 273	14.8	74.0	150351	18	US-10-322-281-453	Sequence 453, App	c 346	14.4	72.0	5139	17	US-10-242-515-3802	Sequence 3802, Ap
c 274	14.8	74.0	154504	18	US-10-322-696-67	Sequence 67, Appl	c 347	14.4	72.0	5404	15	US-10-146-473A-4	Sequence 4, Appli
c 275	14.8	74.0	171936	15	US-10-265-071-24	Sequence 24, Appl	c 348	14.4	72.0	5583	17	US-10-276-774-474	Sequence 474, App
c 276	14.8	74.0	171936	15	US-10-025-966A-24	Sequence 24, Appl	c 349	14.4	72.0	5654	10	US-09-854-867-41	Sequence 41, Appl
c 277	14.8	74.0	182328	17	US-10-235-152A-47	Sequence 47, Appl	c 350	14.4	72.0	7554	17	US-10-313-203-1	Sequence 1, Appli
c 278	14.8	74.0	211357	13	US-10-087-192-529	Sequence 529, App	c 351	14.4	72.0	19000	14	US-10-109-084-3	Sequence 3, Appli
c 279	14.8	74.0	233528	18	US-10-719-993-6856	Sequence 6856, Ap	c 352	14.4	72.0	29111	13	US-10-087-192-865	Sequence 865, App
c 280	14.8	74.0	235070	13	US-10-087-192-1990	Sequence 1990, Ap	c 353	14.4	72.0	100608	18	US-10-417-375-108	Sequence 108, App
c 281	14.8	74.0	254366	10	US-09-822-871-3	Sequence 3, Appli	c 354	14.4	72.0	115284	13	US-10-087-192-673	Sequence 673, App
c 282	14.8	74.0	254366	17	US-10-673-885-3	Sequence 3, Appli	c 355	14.4	72.0	121434	18	US-10-303-165-11	Sequence 11, Appl
c 283	14.8	74.0	261638	18	US-10-719-993-6882	Sequence 6882, Ap	c 356	14.4	72.0	225983	14	US-10-175-523-57	Sequence 57, Appl
c 284	14.8	74.0	296405	13	US-10-087-192-1036	Sequence 1036, Ap	c 357	14.4	72.0	289190	18	US-10-322-281-115	Sequence 115, App
c 285	14.8	74.0	310122	18	US-10-417-375-1	Sequence 1, Appli	c 358	14.4	72.0	392112	17	US-10-812-232-3	Sequence 3, Appli
c 286	14.8	74.0	344548	13	US-10-087-192-334	Sequence 334, App	c 359	14.4	72.0	1830121	18	US-10-329-670-1	Sequence 1, Appli
c 287	14.8	74.0	510510	19	US-10-741-600-17606	Sequence 17606, A	c 360	14.4	72.0	1830121	18	US-10-158-865-1	Sequence 1, Appli
c 288	14.8	74.0	659158	9	US-09-771-208-20	Sequence 20, Appl	c 361	14.2	71.0	25	14	US-10-215-112-12675	Sequence 12675, A
c 289	14.4	72.0	25	19	US-10-719-900-360387	Sequence 360387,	c 362	14.2	71.0	25	14	US-10-215-112-12676	Sequence 12676, A
c 290	14.4	72.0	25	19	US-10-719-900-418253	Sequence 418253,	c 363	14.2	71.0	25	19	US-10-719-900-383016	Sequence 383016,
c 291	14.4	72.0	264	17	US-10-242-535A-11437	Sequence 11437, A	c 364	14.2	71.0	25	19	US-10-719-900-668526	Sequence 668526,
c 292	14.4	72.0	268	17	US-10-085-783A-11437	Sequence 11437, A	c 365	14.2	71.0	25	19	US-10-719-900-889274	Sequence 889274,
c 293	14.4	72.0	278	18	US-10-425-115-58366	Sequence 58366, A	c 366	14.2	71.0	135	17	US-10-242-535A-28620	Sequence 28620, A
c 294	14.4	72.0	316	9	US-09-964-824A-400	Sequence 400, App	c 367	14.2	71.0	135	17	US-10-085-783A-28620	Sequence 28620, A
c 295	14.4	72.0	337	17	US-10-242-535A-2695	Sequence 2695, Ap	c 368	14.2	71.0	165	16	US-10-029-386-16744	Sequence 16744, A
c 296	14.4	72.0	337	17	US-10-242-535A-11019	Sequence 11019, A	c 369	14.2	71.0	255	11	US-09-987-899-3171	Sequence 3171, Ap
c 297	14.4	72.0	337	17	US-10-085-783A-2695	Sequence 2695, Ap	c 370	14.2	71.0	258	9	US-09-867-701-3559	Sequence 3559, Ap
c 298	14.4	72.0	337	17	US-10-085-783A-11019	Sequence 11019, A	c 371	14.2	71.0	261	9	US-09-728-444-1027	Sequence 1027, Ap
c 299	14.4	72.0	341	18	US-10-425-115-17480	Sequence 17480, A	c 372	14.2	71.0	273	18	US-10-437-963-49845	Sequence 49845, A
c 300	14.4	72.0	362	9	US-09-728-444-659	Sequence 659, App	c 373	14.2	71.0	313	18	US-10-425-115-52420	Sequence 52420, A
c 301	14.4	72.0	407	18	US-10-357-930-30894	Sequence 9693, App	c 374	14.2	71.0	322	9	US-09-783-590-7028	Sequence 7028, Ap
c 302	14.4	72.0	453	18	US-10-357-930-30894	Sequence 30894, A	c 375	14.2	71.0	334	18	US-10-425-115-43248	Sequence 43248, A
c 303	14.4	72.0	453	18	US-10-357-930-39863	Sequence 39863, A	c 376	14.2	71.0	356	18	US-10-425-115-17018	Sequence 17018, A

377	14.2	71.0	358	17	US-10-424-599-73392	Sequence 73392, A	C 450	14.2	71.0	624	19	US-10-486-924-123	Sequence 123, App
378	14.2	71.0	394	17	US-10-424-599-127495	Sequence 127495, A	C 451	14.2	71.0	627	18	US-10-767-701-26207	Sequence 26207, A
379	14.2	71.0	410	13	US-10-027-632-279400	Sequence 279400, A	C 452	14.2	71.0	651	18	US-10-767-795-3031	Sequence 3031, Ap
380	14.2	71.0	410	13	US-10-027-632-279400	Sequence 279400, A	C 453	14.2	71.0	655	13	US-10-027-632-105158	Sequence 105158, A
381	14.2	71.0	422	18	US-10-027-632-63067	Sequence 63067, A	C 454	14.2	71.0	655	17	US-10-027-632-105158	Sequence 105158, A
382	14.2	71.0	425	13	US-10-027-633-1258	Sequence 1258, Ap	C 455	14.2	71.0	655	18	US-10-653-047-7387	Sequence 7387, Ap
383	14.2	71.0	425	17	US-10-027-633-1258	Sequence 1258, Ap	C 456	14.2	71.0	663	13	US-10-027-632-186269	Sequence 186269, A
384	14.2	71.0	426	13	US-10-027-632-307040	Sequence 307040, A	C 457	14.2	71.0	663	17	US-10-027-632-186269	Sequence 186269, A
385	14.2	71.0	426	17	US-10-027-632-307040	Sequence 307040, A	C 458	14.2	71.0	666	18	US-10-653-047-1896	Sequence 1896, Ap
386	14.2	71.0	440	17	US-10-424-599-95178	Sequence 95178, A	C 459	14.2	71.0	673	13	US-10-027-632-193897	Sequence 193897, A
387	14.2	71.0	457	19	US-10-696-639-2339	Sequence 2339, App	C 460	14.2	71.0	673	13	US-10-027-632-193897	Sequence 193897, A
388	14.2	71.0	460	17	US-10-424-599-103740	Sequence 103740, A	C 461	14.2	71.0	681	18	US-10-437-963-16438	Sequence 16438, A
389	14.2	71.0	480	18	US-10-437-963-37647	Sequence 37647, A	C 462	14.2	71.0	721	15	US-10-106-698-1888	Sequence 1888, Ap
390	14.2	71.0	481	17	US-10-424-599-47938	Sequence 47938, A	C 463	14.2	71.0	728	18	US-10-425-115-130149	Sequence 130149, A
391	14.2	71.0	499	18	US-10-653-047-2365	Sequence 2365, Ap	C 464	14.2	71.0	732	13	US-10-027-632-154896	Sequence 154896, A
392	14.2	71.0	511	9	US-09-864-761-12487	Sequence 12487, A	C 465	14.2	71.0	732	17	US-10-027-632-154896	Sequence 154896, A
393	14.2	71.0	514	9	US-09-989-919-61	Sequence 919-61, Appl	C 466	14.2	71.0	732	18	US-10-402-339-3	Sequence 3, Appl
394	14.2	71.0	515	11	US-09-864-408A-1497	Sequence 1497, Ap	C 467	14.2	71.0	732	18	US-10-402-339-7	Sequence 7, Appl
395	14.2	71.0	518	13	US-10-027-632-5108	Sequence 5108, Ap	C 468	14.2	71.0	736	13	US-10-027-632-31580	Sequence 31580, A
396	14.2	71.0	518	13	US-10-027-632-5109	Sequence 5109, Ap	C 469	14.2	71.0	736	13	US-10-027-632-31581	Sequence 31581, A
397	14.2	71.0	518	17	US-10-027-632-5109	Sequence 5109, Ap	C 470	14.2	71.0	736	17	US-10-027-632-31580	Sequence 31580, A
398	14.2	71.0	518	17	US-10-027-632-5109	Sequence 5109, Ap	C 471	14.2	71.0	736	17	US-10-027-632-31581	Sequence 31581, A
399	14.2	71.0	528	13	US-10-027-632-285077	Sequence 285077, A	C 472	14.2	71.0	738	13	US-10-027-632-14480	Sequence 14480, A
400	14.2	71.0	528	17	US-10-027-632-285077	Sequence 285077, A	C 473	14.2	71.0	738	17	US-10-027-632-14480	Sequence 14480, A
401	14.2	71.0	537	13	US-10-027-632-279401	Sequence 279401, A	C 474	14.2	71.0	744	13	US-10-027-632-144408	Sequence 144408, A
402	14.2	71.0	537	17	US-10-027-632-279401	Sequence 279401, A	C 475	14.2	71.0	744	13	US-10-027-632-144408	Sequence 144408, A
403	14.2	71.0	545	9	US-10-764-869-1287	Sequence 1287, Ap	C 476	14.2	71.0	744	17	US-10-027-632-144408	Sequence 144408, A
404	14.2	71.0	545	14	US-10-091-504-1287	Sequence 1287, Ap	C 477	14.2	71.0	744	17	US-10-027-632-144409	Sequence 144409, A
405	14.2	71.0	545	17	US-10-027-632-1287	Sequence 1287, Ap	C 478	14.2	71.0	750	9	US-09-912-787-7	Sequence 7, Appl
406	14.2	71.0	549	13	US-10-027-633-65839	Sequence 65839, A	C 479	14.2	71.0	750	9	US-09-912-787-66	Sequence 66, Appl
407	14.2	71.0	549	13	US-10-027-633-65840	Sequence 65840, A	C 480	14.2	71.0	750	13	US-10-014-326-52	Sequence 52, Appl
408	14.2	71.0	549	13	US-10-027-633-65841	Sequence 65841, A	C 481	14.2	71.0	750	13	US-10-014-326-71	Sequence 71, Appl
409	14.2	71.0	549	13	US-10-027-633-65842	Sequence 65842, A	C 482	14.2	71.0	762	18	US-10-767-701-5190	Sequence 5190, Ap
410	14.2	71.0	549	17	US-10-027-632-65839	Sequence 65839, A	C 483	14.2	71.0	764	18	US-10-425-115-135449	Sequence 135449, A
411	14.2	71.0	549	17	US-10-027-632-65840	Sequence 65840, A	C 484	14.2	71.0	774	18	US-10-767-701-7513	Sequence 7513, Ap
412	14.2	71.0	549	17	US-10-027-632-65841	Sequence 65841, A	C 485	14.2	71.0	782	18	US-10-425-115-21489	Sequence 21489, A
413	14.2	71.0	549	17	US-10-027-632-65842	Sequence 65842, A	C 486	14.2	71.0	789	13	US-10-027-632-174033	Sequence 174033, A
414	14.2	71.0	551	16	US-10-029-386-9888	Sequence 9888, Ap	C 487	14.2	71.0	789	13	US-10-027-632-174034	Sequence 174034, A
415	14.2	71.0	557	19	US-10-696-639-238	Sequence 238, App	C 488	14.2	71.0	789	17	US-10-027-632-174033	Sequence 174033, A
416	14.2	71.0	566	13	US-10-027-632-48952	Sequence 48952, A	C 489	14.2	71.0	789	17	US-10-027-632-174034	Sequence 174034, A
417	14.2	71.0	566	13	US-10-027-632-48952	Sequence 48952, A	C 490	14.2	71.0	791	13	US-10-027-632-153438	Sequence 153438, A
418	14.2	71.0	569	13	US-10-027-632-249495	Sequence 249495, A	C 491	14.2	71.0	791	17	US-10-027-632-153438	Sequence 153438, A
419	14.2	71.0	569	17	US-10-027-632-249495	Sequence 249495, A	C 492	14.2	71.0	801	18	US-10-767-701-9882	Sequence 9882, Ap
420	14.2	71.0	573	16	US-10-029-386-1388	Sequence 1388, Ap	C 493	14.2	71.0	807	13	US-10-027-632-148300	Sequence 148300, A
421	14.2	71.0	576	16	US-10-029-386-9690	Sequence 9690, Ap	C 494	14.2	71.0	807	13	US-10-027-632-148301	Sequence 148301, A
422	14.2	71.0	579	18	US-10-437-963-101860	Sequence 101860, A	C 495	14.2	71.0	807	17	US-10-027-632-148300	Sequence 148300, A
423	14.2	71.0	580	10	US-09-918-995-26770	Sequence 26770, A	C 496	14.2	71.0	807	17	US-10-027-632-148301	Sequence 148301, A
424	14.2	71.0	581	13	US-10-027-632-60749	Sequence 60749, A	C 497	14.2	71.0	807	17	US-10-425-114-18341	Sequence 18341, A
425	14.2	71.0	581	13	US-10-027-632-107617	Sequence 107617, A	C 498	14.2	71.0	829	13	US-10-027-632-169165	Sequence 169165, A
426	14.2	71.0	581	13	US-10-027-632-309615	Sequence 309615, A	C 499	14.2	71.0	829	17	US-10-027-632-169165	Sequence 169165, A
427	14.2	71.0	581	17	US-10-027-632-60749	Sequence 60749, A	C 500	14.2	71.0	837	13	US-10-027-632-164807	Sequence 164807, A
428	14.2	71.0	581	17	US-10-027-632-107617	Sequence 107617, A							
429	14.2	71.0	581	17	US-10-027-632-309615	Sequence 309615, A							
430	14.2	71.0	597	16	US-10-029-386-3044	Sequence 3044, Ap							
431	14.2	71.0	598	9	US-09-796-692-4093	Sequence 4093, Ap							
432	14.2	71.0	598	14	US-10-040-862-4093	Sequence 4093, Ap							
433	14.2	71.0	598	17	US-10-057-475B-4093	Sequence 4093, Ap							
434	14.2	71.0	598	17	US-10-154-884B-4093	Sequence 4093, Ap							
435	14.2	71.0	598	18	US-10-764-324-4093	Sequence 4093, Ap							
436	14.2	71.0	607	13	US-10-027-632-261185	Sequence 261185, A							
437	14.2	71.0	607	17	US-10-027-632-261185	Sequence 261185, A							
438	14.2	71.0	621	19	US-10-486-924-107	Sequence 107, App							
439	14.2	71.0	621	19	US-10-486-924-110	Sequence 110, App							
440	14.2	71.0	621	19	US-10-486-924-127	Sequence 127, App							
441	14.2	71.0	621	19	US-10-486-924-129	Sequence 129, App							
442	14.2	71.0	621	19	US-10-486-924-131	Sequence 131, App							
443	14.2	71.0	621	19	US-10-486-924-133	Sequence 133, App							
444	14.2	71.0	621	19	US-10-486-924-135	Sequence 135, App							
445	14.2	71.0	621	19	US-10-486-924-137	Sequence 137, App							
446	14.2	71.0	621	19	US-10-486-924-139	Sequence 139, App							
447	14.2	71.0	621	19	US-10-486-924-141	Sequence 141, App							
448	14.2	71.0	621	19	US-10-486-924-143	Sequence 143, App							
449	14.2	71.0	621	19	US-10-486-924-145	Sequence 145, App							

## ALIGNMENTS

## RESULT 1

US-10-672-399-10

; Sequence 10, Application US/10672399

; Publication No. US20050003368A1

; GENERAL INFORMATION:

; APPLICANT: University of Alberta

; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

; FILE REFERENCE: A894635US

; CURRENT APPLICATION NUMBER: US/10/672,399

; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401

; PRIOR FILING DATE: 2003-05-22

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 10

; LENGTH: 20

; TYPE: DNA

```
; ORGANISM: Homo sapiens
US-10-672-399-10
    Query Match      100.0%; Score 20; DB 18; Length 20;
    Best Local Similarity 100.0%; Pred. No. 1.5; Indels 0; Gaps 0;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||

RESULT 2
US-10-672-399-3/c
; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-3
    Query Match      100.0%; Score 20; DB 18; Length 1071;
    Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 966 AGGGCGTCTCTGAGTAGCAG 947
    |||||

RESULT 3
US-10-672-399-5/c
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5
    Query Match      100.0%; Score 20; DB 18; Length 1083;
    Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 1025 AGGGCGTCTCTGAGTAGCAG 1006
    |||||

RESULT 4
US-10-672-399-1/c
; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1
    Query Match      100.0%; Score 20; DB 18; Length 1737;
    Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGCGTCTCTGAGTAGCAG 20
    |||||
Db 1099 AGGGCGTCTCTGAGTAGCAG 1080
    |||||

RESULT 5
US-10-042-523-1/c
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-042-523-1
```

Query Match 100.0%; Score 20; DB 13; Length 2116;  
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0; Gaps 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
|||||  
Db 1134 AGGCGCTCTCTGAGTAGCAG 1115

RESULT 6  
US-09-902-939-1/c  
; Sequence 1, Application US/09902939  
; Publication No. US20030087850A1  
; GENERAL INFORMATION:  
; APPLICANT: Philip DeHazy  
; APPLICANT: William Chen  
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME  
; FILE REFERENCE: 2055/0H020-US0  
; CURRENT APPLICATION NUMBER: US/09/902,939  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1752  
; TYPE: DNA  
; ORGANISM: mouse  
US-09-902-939-1

Query Match 87.0%; Score 17.4; DB 10; Length 1752;  
Best Local Similarity 94.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCA 19  
|||||  
Db 1114 AGGCGCTCTCCGAGTAGCA 1096

RESULT 7  
US-10-425-115-88239  
; Sequence 88239, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 88239  
; LENGTH: 188  
; TYPE: DNA  
; ORGANISM: zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(188)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_180473C.1  
US-10-425-115-88239

Query Match 84.0%; Score 16.8; DB 18; Length 188;  
Best Local Similarity 90.0%; Pred. No. 63;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
|||||  
Db 99 ATGCGCTTCTGAGTAGCAG 118

RESULT 8  
US-10-029-386-22234/c  
; Sequence 22234, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 22234  
; LENGTH: 239  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR19.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EST HUMAN HIT: BG618999.1, EVALUE 4.00e-47  
; OTHER INFORMATION: SWISSPROT HIT: P01026, EVALUE 3.90e+00  
; OTHER INFORMATION: NT HIT: gill1425973, EVALUE 0.00e+00  
US-10-029-386-22234

Query Match 84.0%; Score 16.8; DB 16; Length 239;  
Best Local Similarity 90.0%; Pred. No. 63;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
|||||  
Db 113 AGGCATCTCTGAGTTGCAG 94

RESULT 9  
US-10-029-386-8534/c  
; Sequence 8534, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 8534  
; LENGTH: 595  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR19.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: SWISSPROT HIT: Q9TC96, EVALUE 1.10e+00  
; OTHER INFORMATION: NT HIT: gill1425973, EVALUE 0.00e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: BF368912.1, EVALUE 8.00e-66  
US-10-029-386-8534

Query Match 84.0%; Score 16.8; DB 16; Length 595;  
Best Local Similarity 90.0%; Pred. No. 60;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCAG 20  
|||||  
Db 422 AGGCATCTCTGAGTTGCAG 403

```

RESULT 10
US-10-322-281-223
; Sequence 223, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 223
; LENGTH: 80105
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(80105)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-223

```

```

Query Match      84.0%; Score 16.8; DB 18; Length 80105;
Best Local Similarity 90.0%; Pred. No. 47;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGGGCGTCTCTGAGTAGCAG 20
      |||||
Db      32700 AGGGCTACTCTGAGTAGCAG 32719
      |||||

```

RESULT 11

US-10-076-555-218

Sequence 218, Application US/10076555

Publication No. US20030065156A1

GENERAL INFORMATION:

APPLICANT: Williams, Lewis T.

APPLICANT: Escobedo, Jaime

APPLICANT: Innis, Michael A.

APPLICANT: Garcia, Pablo Dominguez

APPLICANT: Sudduth-Klinger, Julie

APPLICANT: Reinhard, Christoph

APPLICANT: Giese, Klaus

APPLICANT: Randazzo, Filippo

APPLICANT: Kennedy, Giulia C.

APPLICANT: Pot, David

APPLICANT: Kassam, Altaf

APPLICANT: Lamson, George

APPLICANT: Drmanac, Radoje

APPLICANT: Crkvenjakov, Radomir

APPLICANT: Dickson, Mark

APPLICANT: Drmanac, Snezana

APPLICANT: Labat, Ivan

APPLICANT: Leshkowitz, Dena

APPLICANT: Kita, David

APPLICANT: Garcia, Veronica

APPLICANT: Jones, Lee William

APPLICANT: Stache-Crain, Birgit

TITLE OF INVENTION: Novel Human Genes and Gene Expression

TITLE OF INVENTION: Products I

FILE REFERENCE: 2300-1480

CURRENT APPLICATION NUMBER: US/10/076.555

CURRENT FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/217,471

PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-21

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/068,755

PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-23

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/080,664

PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,234

PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-21

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,877

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 844
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-076-555-218

Query Match      82.0%; Score 16.4; DB 14; Length 300;
Best Local Similarity 94.4%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0

Qy 3 GGCCTCTCTGAGTAGCAG 20
   |||||
Db 222 GGCCTCTCTGAGTAGCAG 239

RESULT 12
US-10-076-555-842
; Sequence 842, Application US/10076555
; Publication No. US20030065156A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Lechkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Novel Human Genes and Gene Expression
; TITLE OF INVENTION: Products I
; FILE REFERENCE: 2300-1480
; CURRENT APPLICATION NUMBER: US/10/076,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/217,473
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/068,751
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,233
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,877
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 844
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 842
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) .. (452)
; OTHER INFORMATION: n = A,T,C or G
; US-10-076-555-842

Query Match      82.0%; Score 16.4; DB 14; Length 452;

```

```
Best Local Similarity 94.4%; Pred. No. 98;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
Db 222 GCGCTCTCTGAGCAGCAG 239

RESULT 13
US-10-260-238-3940/c
; Sequence 3940, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaoki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 3940
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-260-238-3940

Query Match 82.0%; Score 16.4; DB 17; Length 748;
Best Local Similarity 94.4%; Pred. No. 96;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
Db 235 GCGCTCTCCGAGTAGCAG 218

RESULT 14
US-10-104-047-1094
; Sequence 1094, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1094
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1094

Query Match 82.0%; Score 16.4; DB 17; Length 1611;
Best Local Similarity 94.4%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
Db 636 GCGCTCTCTGAGCAGCAG 653

RESULT 15
US-09-925-301-256
; Sequence 256, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 256
; LENGTH: 1992
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: {558}
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-256

Query Match 82.0%; Score 16.4; DB 9; Length 1992;
Best Local Similarity 94.4%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
Db 631 GCGCTCTCTGAGCAGCAG 648

RESULT 16
US-10-264-237-795/c
; Sequence 795, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 795
; LENGTH: 2665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-795

Query Match 82.0%; Score 16.4; DB 17; Length 2665;
Best Local Similarity 94.4%; Pred. No. 90;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
Db 1169 GCGCTCTCTGAGCAGCAG 1152

RESULT 17
```



```
US-10-723-860-6203/c
; Sequence 6203, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05982.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6203
; LENGTH: 4560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-6203

Query Match      82.0%; Score 16.4; DB 18; Length 4560;
Best Local Similarity 94.4%; Pred. No. 87;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCGCTCTCTGAGTAGCAG 20
|||||
Db 2727 GCGCTCTCTGAGTAGCAG 2710
|||||

RESULT 18
US-09-918-995-28588/c
; Sequence 28588, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28588
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28588

Query Match      79.0%; Score 15.8; DB 10; Length 458;
Best Local Similarity 89.5%; Pred. No. 28+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGCGCTCTCTGAGTAGCA 19
|||||
Db 253 AGGGCGCTCTCTGAGTAGCA 235
|||||

RESULT 19
US-09-760-008A-2
; Sequence 2, Application US/09760008A
; Patent No. US20020004483A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
US-09-904-196B-2
; Sequence 2, Application US/09904196B
; Publication No. US20030064922A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/904,196B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/760,008
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-904-196B-2
```

```
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/760,008A
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-760-008A-2

Query Match      79.0%; Score 15.8; DB 9; Length 525;
Best Local Similarity 89.5%; Pred. No. 28+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20
|||||
Db 173 GGGCGCTCTGAGTAGCTG 191
|||||

RESULT 20
US-09-904-196B-2
; Sequence 2, Application US/09904196B
; Publication No. US20030064922A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/09/904,196B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US/09/760,008
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-904-196B-2
```

```
Query Match          79.0%; Score 15.8; DB 10; Length 525;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 173 GGGCGCCTCTGAGTAGCTG 191

RESULT 21
US-10-192-294-2
; Sequence 2, Application US/10192294
; Publication No. US20030118612A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, Torben Lauesgaard
; TITLE OF INVENTION: G-CSF Conjugates
; FILE REFERENCE: 0258us310
; CURRENT APPLICATION NUMBER: US/10/192,294
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 09/904,196
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 00447
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: DK PA 2002 00708
; PRIOR FILING DATE: 2002-05-08
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: DNA sequence encoding hG-CSF, with codon usage for E. coli
US-10-192-294-2

Query Match          79.0%; Score 15.8; DB 15; Length 525;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 173 GGGCGCCTCTGAGTAGCTG 191

RESULT 22
US-10-318-966-2
; Sequence 2, Application US/10318966
; Publication No. US20030158375A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; APPLICANT: SCHAMBYE, HANS THALSGARD
; TITLE OF INVENTION: G-CSF POLYPEPTIDES AND CONJUGATES
; FILE REFERENCE: 0208us620
; CURRENT APPLICATION NUMBER: US/10/318,966
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/904,196
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/760,008
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: 000341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-10-318-966-2

Query Match          79.0%; Score 15.8; DB 18; Length 525;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 173 GGGCGCCTCTGAGTAGCTG 191

RESULT 24
US-10-003-496-5
; Sequence 5, Application US/10003496
; Publication No. US20020142964A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen ApS
```

```
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-10-318-966-2
```

```
Query Match          79.0%; Score 15.8; DB 16; Length 525;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 173 GGGCGCCTCTGAGTAGCTG 191
```

```
RESULT 23
US-10-705-745-2
; Sequence 2, Application US/10705745
; Publication No. US20040241806A1
; GENERAL INFORMATION:
; APPLICANT: NISSEN, TORBEN LAUESGAARD
; APPLICANT: ANDERSEN, KIM VILBOUR
; APPLICANT: HANSEN, CHRISTIAN KARSTEN
; APPLICANT: MIKKELSEN, JAN MOLLER
; TITLE OF INVENTION: G-CSF CONJUGATES
; FILE REFERENCE: 31-000700US
; CURRENT APPLICATION NUMBER: US/10/705,745
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US/09/760,008
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/176,376
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/189,506
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/215,644
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 00024
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: DK PA 2000 00341
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: DK PA 2000 00943
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-10-705-745-2
```

```
Query Match          79.0%; Score 15.8; DB 18; Length 525;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 173 GGGCGCCTCTGAGTAGCTG 191
```

```
RESULT 24
US-10-003-496-5
; Sequence 5, Application US/10003496
; Publication No. US20020142964A1
; GENERAL INFORMATION:
; APPLICANT: Maxygen ApS
```

```

; APPLICANT: Maxygen Holdings Ltd.
; TITLE OF INVENTION: Single-Chain Polypeptides
; FILE REFERENCE: 0218us210
; CURRENT APPLICATION NUMBER: US/10/003,496
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/245,727
; PRIOR FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA encoding G-CSF copy 2 in the single chain G-CSF dimer
US-10-003-496-5

Query Match          79.0%; Score 15.8; DB 13; Length 531;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GGGCGCTCTCTGAGTAGCAG 20
          ||||| ||||| ||||| |||||
Db      173 GGGCGCTCTGAGTAGCTG 191

RESULT 25
US-10-437-963-27903
; Sequence 27903, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 27903
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32553C.1
US-10-437-963-27903

Query Match          79.0%; Score 15.8; DB 18; Length 600;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  AGGGCGCTCTCTGAGTAGCA 19
          ||||| ||||| ||||| |||||
Db      249 AGGGCGCTCTCTGAGGAGCA 267

RESULT 26
US-10-027-632-142002
; Sequence 142002, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

QY      2  GGGCGCTCTCTGAGTAGCAG 20
          ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTATCAG 420

RESULT 27
US-10-027-632-169596
; Sequence 169596, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169596
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION:
US-10-027-632-169596

Query Match          79.0%; Score 15.8; DB 13; Length 874;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GGGCGCTCTCTGAGTAGCAG 20
          ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTATCAG 420

RESULT 28
US-10-027-632-169596
; Sequence 169596, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169596
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION:
US-10-027-632-169596

Query Match          79.0%; Score 15.8; DB 13; Length 874;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GGGCGCTCTCTGAGTAGCAG 20
          ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTATCAG 420

```

```
RESULT 28
US-10-027-632-142002
; Sequence 142002, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142002
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-142002

Query Match          79.0%; Score 15.8; DB 17; Length 874;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGGCGCTCTCTGAGTAGCA 20
        ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTAGCA 420

RESULT 29
US-10-027-632-169596
; Sequence 169596, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169596
; LENGTH: 874
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-169596

Query Match          79.0%; Score 15.8; DB 17; Length 874;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGGCGCTCTCTGAGTAGCA 20
        ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTAGCA 420
```

```
US-10-027-632-169596

Query Match          79.0%; Score 15.8; DB 17; Length 874;
Best Local Similarity 89.5%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GGGCGCTCTCTGAGTAGCA 20
        ||||| ||||| ||||| |||||
Db      402 GGGCGCTCTGAGTAGCA 420

RESULT 30
US-10-027-632-116504/c
; Sequence 116504, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116504
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-116504

Query Match          79.0%; Score 15.8; DB 13; Length 1113;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 AGGGCGCTCTCTGAGTAGCA 19
        ||||| ||||| ||||| |||||
Db      279 AGGGCGCTCTCTGAGGAGCA 261

RESULT 31
US-10-027-632-116504/c
; Sequence 116504, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 116504  
; LENGTH: 1113  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-116504

Query Match 79.0%; Score 15.8; DB 17; Length 1113;  
Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGCGCTCTCTGAGTAGCA 19  
||||| ||||| ||||| ||||| |||||  
Db 279 AGGCGCTCTCTGAGGACCA 261

RESULT 32  
US-10-424-599-16815/c  
; Sequence 16815, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 16815  
; LENGTH: 1196  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_11518C.1  
US-10-424-599-16815

Query Match 79.0%; Score 15.8; DB 17; Length 1196;  
Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 401 GGGCGTCTCGGAGTAGCTG 383

RESULT 33  
US-10-437-963-34349/c  
; Sequence 34349, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 34349  
; LENGTH: 1326  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_38372C.1  
US-10-437-963-34349

Query Match 79.0%; Score 15.8; DB 18; Length 1326;  
Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 304 GGGCTTCTCTGAGTATCAG 286

RESULT 34  
US-10-424-599-110517  
; Sequence 110517, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 110517  
; LENGTH: 1938  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_70808C.1  
US-10-424-599-110517

Query Match 79.0%; Score 15.8; DB 17; Length 1938;  
Best Local Similarity 89.5%; Pred. No. 1.9e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGCAG 20  
||||| ||||| ||||| ||||| |||||  
Db 1040 GGGCGTCTCTAGTTGCAG 1058

RESULT 35  
US-10-437-963-67253/c  
; Sequence 67253, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 67253  
; LENGTH: 2196  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68127C.1
US-10-437-963-67253

Query Match          79.0%; Score 15.8; DB 18; Length 2196;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGCGTCTCTGAGTAGCA 19
    ||||| ||||| ||||| |||||
Db 2061 AGGCATCTCTGAGAGCA 2043

RESULT 36
US-10-027-632-111979/c
; Sequence 111979, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111979
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-111979

Query Match          79.0%; Score 15.8; DB 13; Length 2597;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 2018 GGGCGTTTCTGAGGAGCAG 2000

RESULT 37
US-10-027-632-111979/c
; Sequence 111979, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 111979
; LENGTH: 2597
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-111979

Query Match          79.0%; Score 15.8; DB 17; Length 2597;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 2018 GGGCGTTTCTGAGGAGCAG 2000

RESULT 38
US-10-160-865-15/c
; Sequence 15, Application US/10160865
; Publication No. US20020169139A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; TITLE OF INVENTION: SINGLE GENE ENCODING AORTIC-SPECIFIC AND STRIATED-SPECIFIC
; FILE REFERENCE: 05433/038001
; CURRENT APPLICATION NUMBER: US/10/160,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/134,250
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: US 08/795,868
; PRIOR FILING DATE: 1997-02-06
; PRIOR APPLICATION NUMBER: US 08/494,577
; PRIOR FILING DATE: 1995-06-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1806)
US-10-160-865-15

Query Match          79.0%; Score 15.8; DB 13; Length 2614;
Best Local Similarity 89.5%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGCAG 20
    ||||| ||||| ||||| |||||
Db 1273 GGGCGTCTCTGAGGAGCAG 1255

RESULT 39
US-09-791-105-1/c
; Sequence 1, Application US/09791105
; Patent No. US200202225A1
; GENERAL INFORMATION:
; APPLICANT: Epidauros Biotechnologie AG
; TITLE OF INVENTION: Method for detecting the presence of at least one single allele of
; FILE REFERENCE: EPI 01/00
; CURRENT APPLICATION NUMBER: US/09/791,105
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: EP 00103844.7
```

QY	DB	SEQUENCE	SCORE	DB 18	LENGTH	MISMATCHES	INDELS	GAPS
2	3918	GGCGCTCTCTGAGTAGCAG 20	79.0%	15.8	5000	0	2	0
<p>Query Match Best Local Similarity 89.5%; Pred. No. 1.8e+02; Matches 17; Conservative 0; Mismatches 2; Indels 2; Gaps 0;</p>								
<p>RESULT 42</p> <p>US-10-491-566-87/c</p> <p>Sequence 87, Application US/10491566 Publication No. US20040249144A1 GENERAL INFORMATION: APPLICANT: OriGene Technologies Inc TITLE OF INVENTION: Regulated Breast Cancer Genes FILE REFERENCE: 3U 103 R1 CURRENT APPLICATION NUMBER: US/10/491.566 CURRENT FILING DATE: 2004-04-05 NUMBER OF SEQ ID NOS: 148 SOFTWARE: PatentIn version 3.1 SEQ ID NO 87 LENGTH: 5866 TYPE: DNA ORGANISM: Homo sapiens FEATURE: NAME/KEY: CDS LOCATION: (1)..(5721) OTHER INFORMATION: US-10-491-566-87</p>								
2	3918	GGCGCTCTCTGAGTAGCAG 20	79.0%	15.8	5000	0	2	0
<p>Query Match Best Local Similarity 89.5%; Pred. No. 1.8e+02; Matches 17; Conservative 0; Mismatches 2; Indels 2; Gaps 0;</p>								
<p>RESULT 43</p> <p>US-09-997-722-124/c</p> <p>Sequence 124, Application US/09997722 Publication No. US20040072154A1 GENERAL INFORMATION: APPLICANT: Morris, David TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER FILE REFERENCE: A-71171/RMS/DCF CURRENT APPLICATION NUMBER: US/09/997.722 CURRENT FILING DATE: 2001-11-30 PRIOR APPLICATION NUMBER: US 09/747,377 PRIOR FILING DATE: 2000-12-22 PRIOR APPLICATION NUMBER: US 09/798,586 PRIOR FILING DATE: 2001-03-02 NUMBER OF SEQ ID NOS: 301 SOFTWARE: PatentIn version 3.1 SEQ ID NO 124 LENGTH: 53795 TYPE: DNA ORGANISM: Homo sapiens US-09-997-722-124</p>								
2	21159	GGCGCTCTCTGAGTAGCAG 20	79.0%	15.8	5000	0	2	0
<p>Query Match Best Local Similarity 89.5%; Pred. No. 1.6e+02; Matches 17; Conservative 0; Mismatches 2; Indels 2; Gaps 0;</p>								
<p>RESULT 44</p> <p>US-10-450-826-68/c</p> <p>Sequence 68, Application US/10450826 Publication No. US20040101818A1 GENERAL INFORMATION: US-10-450-826-68/c</p>								

```

; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathon S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Bistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 73465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. AL109748
US-10-450-826-68

Query Match          79.0%; Score 15.8; DB 18; Length 73465;
Best Local Similarity 89.5%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2  GGGCGTCTCTGAGTAGCAG 20
Db   23585 GAGCGTCTGTGAGTAGCAG 23567

RESULT 45
US-10-148-806-3
; Sequence 3, Application US/10148806
; Publication No. US20030138933A1
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-148-806-3

Query Match          79.0%; Score 15.8; DB 15; Length 114793;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2  GGGCGTCTCTGAGTAGCAG 20
Db   49994 GGGCCTCTCTGAGGACGAG 50012

RESULT 46
US-10-723-860-199
; Sequence 199, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: Methods of diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 199
; LENGTH: 153170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-199

Query Match          79.0%; Score 15.8; DB 18; Length 153170;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1  AGGGCGTCTCTGAGTAGCA 19
Db   119585 AGGGCCTCTCTGAGGACGA 119603

RESULT 47
US-10-425-115-71771
; Sequence 71771, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71771
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165459C.1
US-10-425-115-71771

Query Match          77.0%; Score 15.4; DB 18; Length 397;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1  AGGGCGTCTCTGAGTAG 17
Db   4  AGGGCGTCCCTGAGTAG 20

RESULT 48
US-10-653-047-1507/c
; Sequence 1507, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
```

```

; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: Methods of diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 199
; LENGTH: 153170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-199

Query Match          79.0%; Score 15.8; DB 18; Length 153170;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  1  AGGGCGTCTCTGAGTAGCA 19
Db   119585 AGGGCCTCTCTGAGGACGA 119603

RESULT 47
US-10-425-115-71771
; Sequence 71771, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 71771
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_165459C.1
US-10-425-115-71771

Query Match          77.0%; Score 15.4; DB 18; Length 397;
Best Local Similarity 94.1%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  1  AGGGCGTCTCTGAGTAG 17
Db   4  AGGGCGTCCCTGAGTAG 20

RESULT 48
US-10-653-047-1507/c
; Sequence 1507, Application US/10653047
; Publication No. US20040229367A1
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Jeffrey R. Shuster
; APPLICANT: Sakari Kauppinen
; APPLICANT: Ib Groth Clausen
; APPLICANT: Peter Bjarke Olsen
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 5849.200-US
```



; CURRENT APPLICATION NUMBER: US/10/653,047  
; CURRENT FILING DATE: 2003-08-29  
; PRIOR APPLICATION NUMBER: US/09/533,559  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 09/273,623  
; PRIOR FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 7860  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1507  
; LENGTH: 431  
; TYPE: DNA  
; ORGANISM: Fusarium venenatum  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(431)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-653-047-1507

Query Match 77.0%; Score 15.4; DB 18; Length 431;  
Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCGTCTCTGAGTAGC 18  
DB 103 GGGCGTATCTGAGTAGC 87

## RESULT 49

US-09-560-863-242  
; Sequence 242, Application US/09560863  
; Patent No. US20020110809A1  
; GENERAL INFORMATION:  
; APPLICANT: Nehls, Michael C.  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. US20020110809A1el Human Polynucleotides and the  
; FILE REFERENCE: Lex-0018-USA

; CURRENT APPLICATION NUMBER: US/09/560,863  
; CURRENT FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/132,408  
; PRIOR FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 1008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 242  
; LENGTH: 433  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(433)  
; OTHER INFORMATION: n = A,T,C or G

US-09-560-863-242

Query Match 77.0%; Score 15.4; DB 9; Length 433;  
Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGTCTCTGAGTAGCAG 20  
DB 59 GCCTCTCTGAGTAGCAG 75

## RESULT 50

US-09-783-590-10350/c  
; Sequence 10350, Application US/09783590  
; Patent No. US20020110850A1  
; GENERAL INFORMATION:

; APPLICANT: Dillon, Patrick J.  
; APPLICANT: Haseltine, William A.  
; APPLICANT: Li, Haodong  
; APPLICANT: Rosen, Craig A.  
; APPLICANT: Ruben, Steven M.

; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2  
; FILE REFERENCE: PO-16.2C1  
; CURRENT APPLICATION NUMBER: US/09/783,590  
; CURRENT FILING DATE: 2000-02-15  
; PRIOR APPLICATION NUMBER: 08/420,856  
; PRIOR FILING DATE: 1995-04-12  
; PRIOR APPLICATION NUMBER: 08/346,731  
; PRIOR FILING DATE: 1994-11-21  
; NUMBER OF SEQ ID NOS: 12485  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10350  
; LENGTH: 488  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (9)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (122)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (137)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (188)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (252)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (272)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (314)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (330)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (337)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (340)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (351)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (354)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (363)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (372)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (386)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (394)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (419)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (423)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (426)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature

; LOCATION: (431)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (434)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (440)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (459)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (464)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-783-590-10350

Query Match 77.0%; Score 15.4; DB 9; Length 488;  
Best Local Similarity 94.1%; Pred. No. 3.3e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGGCGTCTCTGAGTAGC 18  
| | | | | | | | | | | | | | | | | | | |  
Db 72 GGGCGTCTCTGAGTAGC 56

Search completed: March 15, 2005, 00:22:50  
Job time : 365 secs